UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

On the inexact scaled gradient projection method

Max Valério Lemes

Goiânia - Goiás

2022

Acknowledgments

Gostaria de agradecer em primeiro lugar a Deus, que sempre proporcionou-me uma vida cheia de bençãos e conquistas. Ao meu orientador Prof. Dr. Orizon Pereira Ferreira, por todos os ensinamentos e direcionamentos, pela infindável paciência e admirável compreensão, além do seu determinante auxílio para a concretização desta tese. Ao meu coorientador Prof. Dr. Glaysdston de Carvalho Bento, pelos ensinamentos e pela enorme contribuição para a realização deste trabalho. Aos professores João Xavier da Cruz Neto, Maicon Marques Alves e Luís Román Lucambio Pérez, por aceitarem participar da minha banca de defesa e pelas valiosas sugestões. Estendo meus agradecimentos à CAPES pela grande ajuda oferecida com a bolsa de estudos de doutorado, à Universidade Federal de Goiás (UFG), pela aceitação no seu programa de doutorado como aluno regular e os conhecimentos adquiridos. Aos professores da UFG pelo apoio e dedicação, em especial aos que integram o grupo de Otimização, e a todos os funcionários da UFG. Finalmente, devo muita gratidão a minha esposa (Luama), irmãos (Adriano e Lívia), mãe (Gleice), avó (Dalva) e a todos os amigos que me acolheram aqui em Goiânia e aos queridos amigos do Piauí; vocês foram muito importantes para mim todos esses anos.

Abstract

In this thesis, we study variational inequalities and generalized vector equilibrium problems.

In Chapter **??**, several results and basic definitions of Riemannian geometry are listed; we present the concept of the monotone vector field in Hadamard manifolds and many of their properties, besides, we introduce the concept of enlargement of a monotone vector field, and we display its properties in a Riemannian context.

In Chapter **??**, an inexact proximal point method for variational inequalities in Hadamard manifolds is introduced, and its convergence properties are studied; see [**?**]. To present our method, we generalize the concept of enlargement of monotone operators, from a linear setting to the Riemannian context. As an application, an inexact proximal point method for constrained optimization problems is obtained.

In Chapter **??**, we present an extragradient algorithm for variational inequality associated with the point-to-set vector field in Hadamard manifolds and study its convergence properties; see [**?**]. In order to present our method, the concept of enlargement of maximal monotone vector fields is used and its lower-semicontinuity is established to obtain the convergence of the method in this new context.

In Chapter **??**, we present a sufficient condition for the existence of a solution to the generalized vector equilibrium problem on Hadamard manifolds using a version of the Knaster-Kuratowski-Mazurkiewicz Lemma; see [**?**]. In particular, the existence of solutions to optimization, vector optimization, Nash equilibria, complementarity, and variational inequality is a special case of the existence result for the generalized vector equilibrium problem.

Keywords: Enlargement of vector fields; inexact proximal; constrained optimization; extragradient algorithm; lower-semicontinuity; vector equilibrium problem; vector optimization; Hadamard manifold.

Contents

List of Figures

List of Algorithms

Introduction

This work is devoted to the study of the *scaled gradient projection (SGP) method with nonmonotone line search* to solve general constrained convex optimization problems as follows

$$\min\{f(x) : \ x \in C\}, \tag{1}$$?{eq:OptP

where C is a closed and convex subset of $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. Denotes by $f^* := \inf_{x\in C} f(x)$ the optimal value of (1) and by Ω^* its solution set, which we will assume to be nonempty unless the contrary is explicitly stated. Problem (1) is a basic issue of constrained optimization, which appears very often in various areas, including finance, machine learning, control theory, and signal processing, see for example [22, 23, 37, 49, 53, 68]. Recent problems considered in most of these areas, the datasets are large or high-dimensional and their solutions need to be approximated quickly with a reasonably accuracy. It is well known that SGP method with nonmonotone line search is among those that are suitable for this task, as will be explained below.

The gradient projection method is what first comes to mind when we start from the ideas of the classic optimization methods in an attempt to deal with problem (1). In fact, this method is one of the oldest known optimization methods to solve (1), the study of its convergence properties goes back to the works of Goldstein [42] and Levitin and Polyak [52]. After these works, several variants of it have appeared over the years, resulting in a vast literature on the subject, including [11, 12, 13, 35, 37, 45, 50, 60, 74]. Additional reference on this subject can be found in the recent review [19] and references therein. Among all the variants of the gradient projection method, the scaled version has been especially considered due to the flexibility provided in efficient implementations of the method; see [15, 5, 18, 20, 21]. In addition, its simplicity and easy implementation has attracted the attention of the scientific community that works on optimization over the years. This method usually uses only first-order derivatives, which makes it very stable from a numerical point of view and therefore quite suitable for solving large-scale optimization problems, see [56, 57, 68, 69]. At each current iteration, SGP method moves along the direction of the negative scaled gradient, and then projects the obtained point onto the constraint set. The current iteration and such projection define a feasible descent direction and a line search in this direction is performed to define the next iteration. In this way, the performance of the method is strongly related to each of the steps we have just mentioned. In fact, the scale matrix and the step size towards the negative scaled gradient are freely selected in order to improve the performance of SGP method but without increasing the cost of each iteration. Strategies for choosing both have their origins in the study

of gradient method for unconstrained optimization, papers addressing this issues include but not limited to [8, 20, 28, 29, 31, 38, 76, 27, 52]. It is worth mentioning that, for suitable choices of the scale matrix and the step size, SGP merges into the well known *spectral gradient method* extensively studied in [15, 14]. More details about selecting step sizes and scale matrices can be found in the recent review [19] and references therein.

In this paper, we are particularly interested in the main stages that make up the SGP method, namely, in the projection calculation and in the line search employed. It is well known that the mostly computational burden of each iteration of the SGP method is in the calculation of the projection. Indeed, the projection calculation requires, at each iteration, the solution of a quadratic problem restricted to the feasible set, which can lead to a substantial increase in the cost per iteration if the number of unknowns is large. For this reason, it may not be justified to carry out exact projections when the iterates are far from the solution of the problem. In order to reduce the computational effort spent on projections, inexact procedures that become more and more accurate when approaching the solution, have been proposed, resulting in more efficient methods; see for exemple [15, 18, 41, 44, 66, 71, 62]. On the other hand, nonmonotone searches can improve the probability of finding an optimal global solution, in addition to potentially improving the speed of convergence of the method as a whole, see for example [26, 59, 70]. The concept of nonmonotone line search, that we will use here as a synonym for inexact line search, have been proposed first in [48], and later a new nonmonotone search was proposed in [75]. After these papers others nonmonotone searches appeared, see for example [3, 55]. In [65], an interesting general framework for nonmonotone line search was proposed, and more recently modifications of it have been presented in [46, 47].

The purpose of the present paper is to present an inexact version of the SGP method, which is inexact in two sense. First, using a version of scheme introduced in [15] and also a variation of the one appeared in [71, Example 1], the inexact projection onto the feasible set is computed allowing an appropriate relative error tolerance. Second, using the inexact conceptual scheme for the line search introduced in [47, 65], a step size is computed to define the next iteration. More specifically, initially we show that the feasible inexact projection of [15] provides greater latitude than the projection of [71, Example 1]. In the first convergence result presented, we show that the SGP method using the projection proposed in [15] preserves the same partial convergence result as the classic method, that is, we prove that every accumulation point of the sequence generated by the SGP method is stationary for problem (1). Then, considering the inexact projection of [71, Example 1], and under mild assumptions, we establish full asymptotic convergence results and some complexity bounds. The presented analysis of the method is done using the general nonmonotone line search scheme introduced in [47]. In this way, the proposed method can be adapted to several line searches and, in particular, will allow obtaining several known versions of the SGP method as particular instances, including [11, 15, 50, 73]. Except for the particular case when we assume that the SGP method employs the nonmonotone line search introduced by [48], all other asymptotic convergence and complexity results are obtained without any assumption of

compactness of the sub-level sets of the objective function. Finally, it is worth mentioning that the complexity results obtained for the SGP method with a general nonmonotone line search are the same as in the classic case when the usual Armijo search is employed, namely, the complexity bound $\mathcal{O}(1/\sqrt{k})$ is unveil for finding ϵ-stationary points for problem (1) and, under convexity on f, the rate to find a ϵ-optimal functional value is $\mathcal{O}(1/k)$.

In Section 1, some notations and basic results used throughout the paper is presented. In particular, Section 2 is devoted to recall the concept of relative feasible inexact projection and some new properties about this concept are presented. Section 3 describes the SGP method with a general nonmonotone line search and some particular instances of it are presented. Partial asymptotic convergence results are presented in Section 3.1. Section **??** presents a full convergence result and iteration-complexity bounds. Some numerical experiments are provided in Section 4. Finally, some concluding remarks are made in Section 6.

Chapter 1

Preliminaries

In this chapter, we introduce some notation and results used throughout our presentation. We also introduce the concept of risk measure that will be used in Chapter 5.

First we consider the index set $\mathbb{N} := \{0, 1, 2, \ldots\}$, the usual inner product $\langle \cdot, \cdot \rangle$ in $\mathbb{R}^n$, and the associated Euclidean norm $\|\cdot\|$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and $C \subseteq \mathbb{R}^n$. The gradient ∇f of f is said to be *Lipschitz continuous* in C with constant $L > 0$ if $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$, for all $x, y \in C$. Combining this definition with the fundamental theorem of calculus, we obtain the following result whose proof can be found in [13, Proposition A.24].

Lemma 1.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and $C \subseteq \mathbb{R}^n$. Assume that ∇f is Lipschitz continuous in C with constant $L > 0$. Then,*

$$f(y) - f(x) - \langle \nabla f(x), y - x \rangle \leq (L/2)\|x - y\|^2,$$

for all $x, y \in C$.

Assume that C is a convex set. The function f is said to be *convex* on C, if

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle,$$

for all $x, y \in C$. We recall that a point $\bar{x} \in C$ is a *stationary point* for problem (1) if

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0, \qquad \forall\, x \in C. \tag{1.1}$$?{eq:Stat

Consequently, if f is a convex function on C, then (1.1) implies that $\bar{x} \in \Omega^*$. We will now present some useful concepts for the analysis of the sequence generated by the scaled gradient method, for more details, see [25]. For that, let D be a $n \times n$ positive definite matrix and $\|\cdot\|_D : \mathbb{R}^n \to \mathbb{R}$ be the norm defined by

$$\|d\|_D := \sqrt{\langle Dd, d \rangle}, \quad \forall d \in \mathbb{R}^n. \tag{1.2}$$?{def:nor

For a fixed constant $\mu \geq 1$, *denote by $\mathcal{D}_\mu$ the set of symmetric positive definite matrices $n \times n$ with all eigenvalues contained in the interval $[\frac{1}{\mu}, \mu]$*. The set $\mathcal{D}_\mu$ is compact. Moreover, for each $D \in \mathcal{D}_\mu$, it follows that D^{-1} also belongs to $\mathcal{D}_\mu$. Furthermore, due to $D \in \mathcal{D}_\mu$, by (1.2), we obtain

$$\frac{1}{\mu}\|d\|^2 \leq \|d\|_D^2 \leq \mu\|d\|^2, \qquad \forall d \in \mathbb{R}^n. \tag{1.3}$$?{eq:pnv}

Let us recall the the concept of sequence quasi-Fejér monotone to a set, introduced in [25].

Definition 1.2. *Let $(y^k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}^n$ and $(D_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}_\mu$. The sequence $(y^k)_{k\in\mathbb{N}}$ is said to be quasi-Fejér monotone to a set $W\subset\mathbb{R}^n$ with respect to $(D_k)_{k\in\mathbb{N}}$ if, there exists a sequence $(\eta_k)_{k\in\mathbb{N}}\subset[0,+\infty)$ such that $\sum_{k\in\mathbb{N}}\eta_k<\infty$ and for all $w\in W$, there exists a sequence $(\epsilon_k)_{k\in\mathbb{N}}\subset[0,+\infty)$ such that $\sum_{k\in\mathbb{N}}\epsilon_k<\infty$, and*

$$\|y_{k+1}-w\|^2_{D_{k+1}}\leq(1+\eta_k)\|y^k-w\|^2_{D_k}+\epsilon_k,$$

for all $k\in\mathbb{N}$.

The following lemma is useful to study the quasi-Fejér monotone sequence, its prove can be found in [61, Lemma 2.2.2].

Lemma 1.3. *Let $(\alpha_k)_{k\in\mathbb{N}}$, $(\eta_k)_{k\in\mathbb{N}}$ and $(\epsilon_k)_{k\in\mathbb{N}}$ be a sequences in $[0,+\infty)$ such that $\sum_{k\in\mathbb{N}}\eta_k<\infty$ and $\sum_{k\in\mathbb{N}}\epsilon_k<\infty$. Assume that $\alpha_{k+1}\leq(1+\eta_k)\alpha_k+\epsilon_k$, for all $k\in\mathbb{N}$. Then, $(\alpha_k)_{k\in\mathbb{N}}$ converges.*

The main property of quasi-Fejér monotone sequences is stated in the following. Its proof can be found in [25, Proposition 3.2 and Theorem 3.3]. For sake of completeness, we include it here.

Theorem 1.4. *Let $(y^k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}^n$ and $(D_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}_\mu$ such that $\lim_{k\to\infty}D_k=\bar{D}$. If $(y^k)_{k\in\mathbb{N}}$ is quasi-Fejér monotone to a nonempty set $W\subset\mathbb{R}^n$ with respect to $(D_k)_{k\in\mathbb{N}}$ then, for each $w\in W$, the sequence $(\|y^k-w\|_{D_k})_{k\in\mathbb{N}}$ converges. Furthermore, $(y^k)_{k\in\mathbb{N}}$ is bounded and, if each cluster point of $(y^k)_{k\in\mathbb{N}}$ belongs to W, then there exists $\bar{y}\in W$ such that $\lim_{k\to\infty}y^k=\bar{y}$.*

Proof. Take $w\in W$ and define the sequence $(\alpha_k)_{k\in\mathbb{N}}$, where $\alpha_k:=\|y^k-w\|_{D_k}$. Since $(y^k)_{k\in\mathbb{N}}$ is quasi-Fejér monotone to W, Lemma 1.3 implies that $(\alpha_k)_{k\in\mathbb{N}}$ converges. Now, by using the first inequality in (1.3), we have $\|y^k-w\|\leq\sqrt{\mu}\alpha_k$, for all $k\in\mathbb{N}$. Thus, $(y^k)_{k\in\mathbb{N}}$ is bounded. To prove the last statement, assume that $\bar{y},\hat{y}\in W$ are cluster points of $(y^k)_{k\in\mathbb{N}}$, and set $(y^{k_i})_{i\in\mathbb{N}}$ and $(y^{k_j})_{j\in\mathbb{N}}$ subsequences of $(y^k)_{k\in\mathbb{N}}$ such that $\lim_{j\to+\infty}y^{k_i}=\bar{y}$ and $\lim_{j\to+\infty}y^{k_j}=\hat{y}$. It follows from the first statement that $(\|y^k-\bar{y}\|_{D_k})_{k\in\mathbb{N}}$ and $(\|y^k-\hat{y}\|_{D_k})_{k\in\mathbb{N}}$ are convergent. Since $\lim_{k\to\infty}D_k=\bar{D}$, we have $\lim_{k\to\infty}\|\bar{y}\|_{D_k}=\|\bar{y}\|_{\bar{D}}$ and $\lim_{k\to\infty}\|\hat{y}\|_{D_k}=\|\hat{y}\|_{\bar{D}}$. Hence, due to

$$\langle y^k,D_k(\bar{y}-\hat{y})\rangle=\frac{1}{2}(\|y^k-\hat{y}\|^2_{D_k}-\|y^k-\bar{y}\|_{D_k}+\|\bar{y}\|_{D_k}-\|\hat{y}\|_{D_k}),$$

for all $k\in\mathbb{N}$, we conclude that the sequence $(\langle y^k,D_k(\bar{y}-\hat{y})\rangle)_{k\in\mathbb{N}}$ converges. Thus, taking into account that $\lim_{j\to+\infty}y^{k_i}=\bar{y}$, $\lim_{j\to+\infty}y^{k_j}=\hat{y}$ and $\lim_{k\to\infty}D_k=\bar{D}$ we obtain that

$$\langle\bar{y},\bar{D}(\bar{y}-\hat{y})\rangle=\lim_{i\to\infty}\langle y^{k_i},D_{k_i}(\bar{y}-\hat{y})\rangle=\lim_{j\to\infty}\langle y^{k_j},D_{k_j}(\bar{y}-\hat{y})\rangle=\langle\hat{y},\bar{D}(\bar{y}-\hat{y})\rangle.$$

Hence, using (1.3), we obtain

$$\frac{1}{\mu}\|\bar{y}-\hat{y}\|^2\leq\|\bar{y}-\hat{y}\|^2_{\bar{D}}=\langle\bar{y},\bar{D}(\bar{y}-\hat{y})\rangle-\langle\hat{y},\bar{D}(\bar{y}-\hat{y})\rangle=0,$$

which implies that $\bar{y}=\hat{y}$. Therefore, due to $(y^k)_{k\in\mathbb{N}}$ be bounded, we conclude that $(y^k)_{k\in\mathbb{N}}$ converges to $\bar{y}$. □

Chapter 2

Relative feasible inexact projections

In this chapter, we recall two concepts of relative feasible inexact projections onto a closed and convex set, and also present some new properties of them which will be used throughout this work. These concepts of inexact projections were introduced seeking to make the subproblem of computing the projections on the feasible set more efficient; see for example [15, 66, 71]. Before presenting the inexact projection concept that we will use, let us first recall the concept of exact projection with respect to a given norm. For that, *throughout this chapter* $D \in \mathcal{D}_\mu$.

Definition 2.1. *The exact projection of the point* $v \in \mathbb{R}^n$ *onto* C *with respect to the norm* $\|\cdot\|_D$*, denoted by* $\mathcal{P}_C^D(v)$*, is defined by*

$$\mathcal{P}_C^D(v) := \arg\min_{z \in C} \|z - v\|_D^2. \tag{2.1}$$?{eq:exac

The next result characterizes the exact projection, its proof can be found in [9, Theorem 3.14].

Lemma 2.2. *Let* $v, w \in \mathbb{R}^n$*. Then,* $w = \mathcal{P}_C^D(v)$ *if and only if* $w \in C$ *and*

$$\langle D(v - w), y - w \rangle \leq 0, \tag{2.2}$$?{eq:exac

for all $y \in C$*.*

Remark 2.3. It follows from Definition 2.1 that $w = \mathcal{P}_C^D(v)$ is the point of C more close to v with respect to the norm $\|\cdot\|_D$. On the other hand, by the Lemma 2.2, w is the unique point of C that for all $y \in C$ the angle θ among the vectors $v - w$ and $y - w$ is an obtuse angle. Figure 2.1 shows this fact considering the plane generated by points y, w and v.

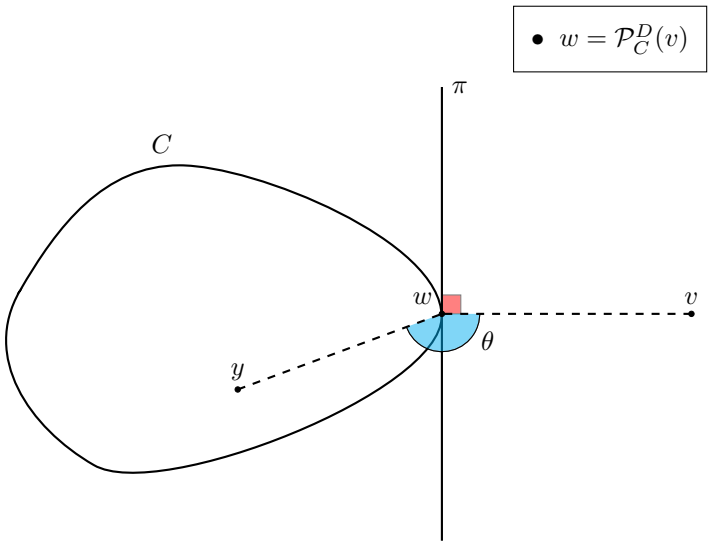


Figure 2.1: Exact projection of the point v onto C.

Remark 2.4. It follows from Lemma 2.2 that $\|\mathcal{P}_C^D(v) - \mathcal{P}_C^D(u)\|_D \leq \|v-u\|_D$. Moreover, since $D \in \mathcal{D}_\mu$, by (1.3), we conclude that $\mathcal{P}_C^D(\cdot)$ is Lipschitz continuous with constant $L = \mu$. Furthermore, if $(D_k)_{k\in\mathbb{N}} \subset \mathcal{D}_\mu$, $\lim_{k\to+\infty} z^k = \bar{z}$, and $\lim_{k\to+\infty} D_k = \bar{D}$, then $\lim_{k\to+\infty} \mathcal{P}_C^{D_k}(z^k) = \mathcal{P}_C^{\bar{D}}(\bar{z})$, see [25, Proposition 4.2].

2.1 The first feasible inexact projection

In the following, we recall the concept of a feasible inexact projection with respect to $\|\cdot\|_D$ relative to a fixed point.

Definition 2.5. *The feasible inexact projection mapping, with respect to the norm $\|\cdot\|_D$, onto C relative to a point $u \in C$ and forcing parameter $\zeta \in (0,1]$, denoted by $\mathcal{P}_{C,\zeta}^D(u,\cdot) : \mathbb{R}^n \rightrightarrows C$, is the set-valued mapping defined as follows*

$$\mathcal{P}_{C,\zeta}^D(u,v) := \left\{ w \in C : \ \|w-v\|_D^2 \leq \zeta \|\mathcal{P}_C^D(v) - v\|_D^2 + (1-\zeta)\|u-v\|_D^2 \right\}. \tag{2.3}$$?{eq:Proj

Each point $w \in \mathcal{P}_{C,\zeta}^D(u,v)$ is called a feasible inexact projection, with respect to the norm $\|\cdot\|_D$, of v onto C relative to u and forcing parameter $\zeta \in (0,1]$.

Remark 2.6. It follows from Definition 2.5 that $\mathcal{P}_{C,\zeta}^D(u,v)$ is a set generated by the intersection of C and a sphere centered in v and radius given by

$$\zeta\|\mathcal{P}_C^D(v) - v\|_D^2 + (1-\zeta)\|u-v\|_D^2.$$

If $\zeta = 1$, then $\mathcal{P}_{C,1}^D(u,v) = \{\mathcal{P}_C^D(v)\}$ is the exact projection of v onto C. However, when ζ is close to zero, the radius of sphere is close to $\|u-v\|_D$. The inexact condition appears when we consider

points that do not minimize the distance from C to v, contrary to property (2.1). This situation is illustrated in Figure 2.2.

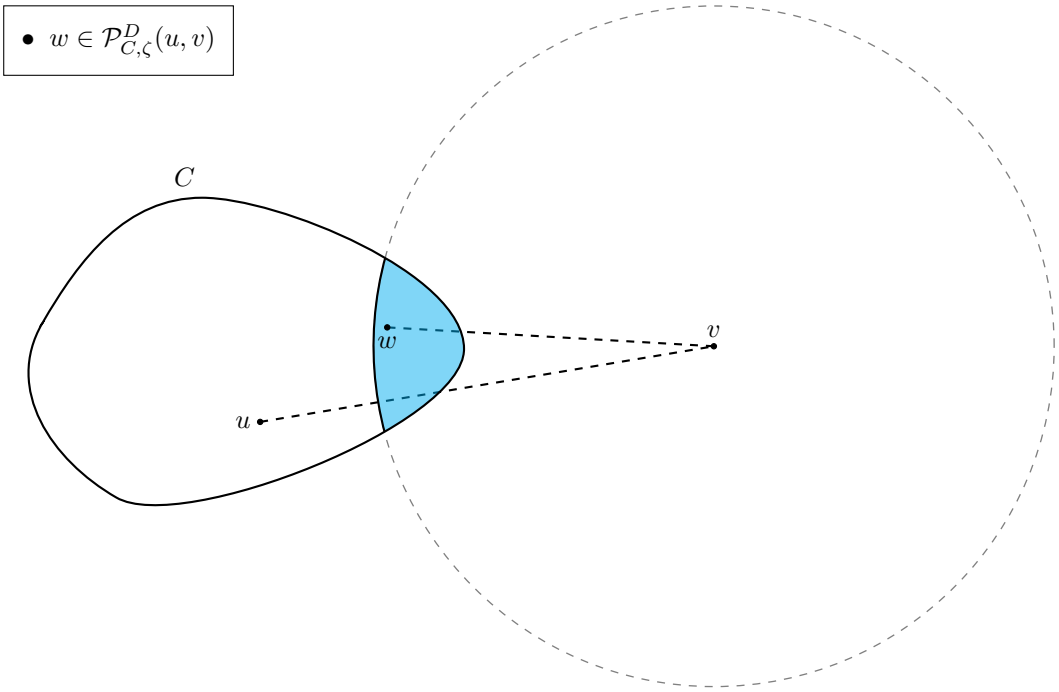


Figure 2.2: Feasible inexact projection of the point v onto C.

In the following, we show that the definition given above is nothing more than a reformulation of the concept of relative feasible inexact projection with respect to $\|\cdot\|_D$ introduced in [15].

Remark 2.7. Let $u \in C$, $v \in \mathbb{R}^n$ and D be an $n \times n$ positive definite matrix. Consider the quadratic function $Q : \mathbb{R}^n \to \mathbb{R}$ defined by

$$Q(z) := (1/2)\,\langle D(z-u), z-u\rangle + \langle D(u-v), z-u\rangle\,.$$

Thus, letting $\|\cdot\|_D$ be the norm defined by (1.2), some algebraic manipulations shows that

$$\|z-v\|_D^2 = 2Q(z) + \|u-v\|_D^2. \tag{2.4}$$?eq:qppq

Hence, (2.4) and (2.1) implies that $\mathcal{P}_C^D(v) = \arg\min_{z\in C} Q(z)$. Let $\zeta \in (0,1]$. Thus, by using (2.4), after some calculations, we can see that the following inexactness condition introduced in [15],

$$w \in C, \qquad Q(w) \le \zeta Q(\mathcal{P}_C^D(v)),$$

is equivalent to find $w \in C$ such that

$$\|w-v\|_D^2 \le \zeta\|\mathcal{P}_C^D(v)-v\|_D^2 + (1-\zeta)\|u-v\|_D^2,$$

which corresponds to condition (2.3) in Definition 2.5.

The concept of feasible inexact projection in Definition 2.5 provides more latitude to the usual concept of exact projection (2.1). The next remark makes this more precise.

Remark 2.8. Let ζ be positive forcing parameter, $C \subset \mathbb{R}^n$ and $u \in C$ be as in Definition 2.5. First of all note that $\mathcal{P}_C^D(v) \in \mathcal{P}_{C,\zeta}^D(u,v)$. Therefore, $\mathcal{P}_{C,\zeta}^D(u,v) \neq \varnothing$, for all $u \in C$ and $v \in \mathbb{R}^n$. Consequently, the set-valued mapping $\mathcal{P}_{C,\zeta}^D(u,\cdot)$ as stated in (2.3) is well-defined. Moreover, for $\zeta = 1$, we have $\mathcal{P}_{C,1}^D(u,v) = \{\mathcal{P}_C^D(v)\}$. In addition, if $\underline{\zeta}$ and $\bar{\zeta}$ are forcing parameters such that $0 < \underline{\zeta} \leq \bar{\zeta} \leq 1$, then $\mathcal{P}_{C,\bar{\zeta}}^D(u,v) \subset \mathcal{P}_{C,\underline{\zeta}}^D(u,v)$.

Lemma 2.9. *Let $v \in \mathbb{R}^n$, $u \in C$ and $w \in \mathcal{P}_{C,\zeta}^D(u,v)$. Then,*

$$\langle D(v-w), y-w\rangle \leq \frac{1}{2}\|w-y\|_D^2 + \frac{1}{2}\left[\zeta\|\mathcal{P}_C^D(v)-v\|_D^2 + (1-\zeta)\|u-v\|_D^2 - \|y-v\|_D^2\right], \qquad y \in C.$$

Proof. Let $y \in C$. Since $2\langle D(v-w), y-w\rangle = \|w-y\|_D^2 + \|w-v\|_D^2 - \|v-y\|_D^2$, using (2.3) we have $2\langle D(v-w), y-w\rangle = \|w-y\|_D^2 + \zeta\|\mathcal{P}_C^D(v)-v\|_D^2 + (1-\zeta)\|u-v\|_D^2 - \|v-y\|_D^2$, which is equivalent to the desired inequality. □

2.2 The second feasible inexact projection

Next, we recall a second concept of relative feasible inexact projection onto a closed convex set, see [2, 30]. The definition is as follows.

Definition 2.10. *The feasible inexact projection mapping, with respect to the norm $\|\cdot\|_D$, onto C relative to $u \in C$ and forcing parameter $\gamma \geq 0$, denoted by $\mathcal{R}_{C,\gamma}^D(u,\cdot) : \mathbb{R}^n \rightrightarrows C$, is the set-valued mapping defined as follows*

$$\mathcal{R}_{C,\gamma}^D(u,v) := \left\{w \in C : \ \langle D(v-w), y-w\rangle \leq \gamma\|w-u\|_D^2, \quad \forall\ y \in C\right\}. \tag{2.5}$$?{eq:Proj

Each point $w \in \mathcal{R}_{C,\gamma}^D(u,v)$ is called a feasible inexact projection, with respect to the norm $\|\cdot\|_D$, of v onto C relative to u and forcing parameter $\gamma \geq 0$.

The concept of feasible inexact projection in Definition 2.10 also provides more latitude to the usual concept of exact projection. Next, we present some remarks about this concept.

Remark 2.11. Let $\gamma \geq 0$ be a forcing parameter, $C \subset \mathbb{R}^n$ and $u \in C$ be as in Definition 2.10. For all $v \in \mathbb{R}^n$, it follows from (2.5) and Lemma 2.2 that $\mathcal{R}_{C,0}^D(u,v) = \{\mathcal{P}_C^D(v)\}$ is the exact projection of v onto C. Moreover, $\mathcal{P}_C^D(v) \in \mathcal{R}_{C,\gamma}^D(u,v)$ concluding that $\mathcal{R}_{C,\gamma}(u,v) \neq \varnothing$, for all $u \in C$ and $v \in \mathbb{R}^n$. Consequently, the set-valued mapping $\mathcal{R}_{C,\gamma}^D(u,\cdot)$ as stated in (2.5) is well-defined.

We show now a geometric interpretation for the inexact projection defined in Definition 2.10.

Remark 2.12. Let $C \subset \mathbb{R}^n$, $u \in C$, $v \in \mathbb{R}^n$, $w \in \mathcal{R}_{C,\gamma}^D(u,v)$ be as stated in Definition 2.10 and $w_t = w + t(v-w)$, with $0 \leq t < 1$. It is easy to see that, for all $y \in C$,

$$\begin{aligned} w_t - w &= t(v-w) \\ v - w_t &= (1-t)(v-w) \\ y - w_t &= y - w - t(v-w). \end{aligned}$$

It follows from (2.5) that $\langle D(v-w), y-w\rangle \leq \gamma\|w-u\|_D^2$, so we have

$$\begin{aligned}\langle D(v-w_t), y-w_t\rangle &= (1-t)\langle D(v-w), y-w\rangle - t(1-t)\|v-w\|_D^2 \\ &\leq (1-t)\left[\gamma\|w-u\|_D^2 - t\|v-w\|_D^2\right].\end{aligned}$$

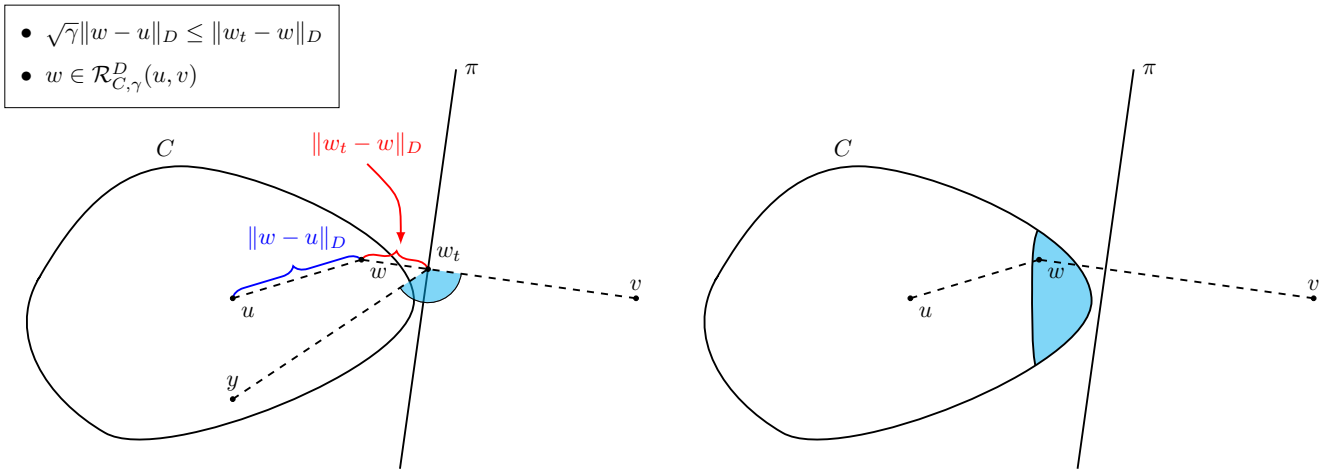
Since $0 \leq t < 1$,

$$\|w_t - w\|_D^2 = t^2\|v-w\|_D^2 \leq t\|v-w\|_D^2,$$

if

$$\gamma\|w-u\|_D^2 \leq \|w_t - w\|_D^2 \tag{2.6}$$?eq:cond

then $\langle D(v-w_t), y-w_t\rangle \leq 0$. In this case the inexact condition appears considering that there is w_t between w and v that satisfies (2.6) and consequently the condition (2.2). This situation is illustrated in Figure 2.3.



(a) Geometric interpretation.

(b) Example of $\mathcal{R}^D_{C,\gamma}(u,v)$.

Figure 2.3: Geometric interpretation of projection $\mathcal{R}^D_{C,\gamma}(u,v)$.

In the following we presents some examples of regions given by inexact projection $\mathcal{R}^D_{C,\gamma}(u,v)$. Note that the projection $\mathcal{R}^D_{C,\gamma}(u,v)$ gets smaller as u gets closer to w.

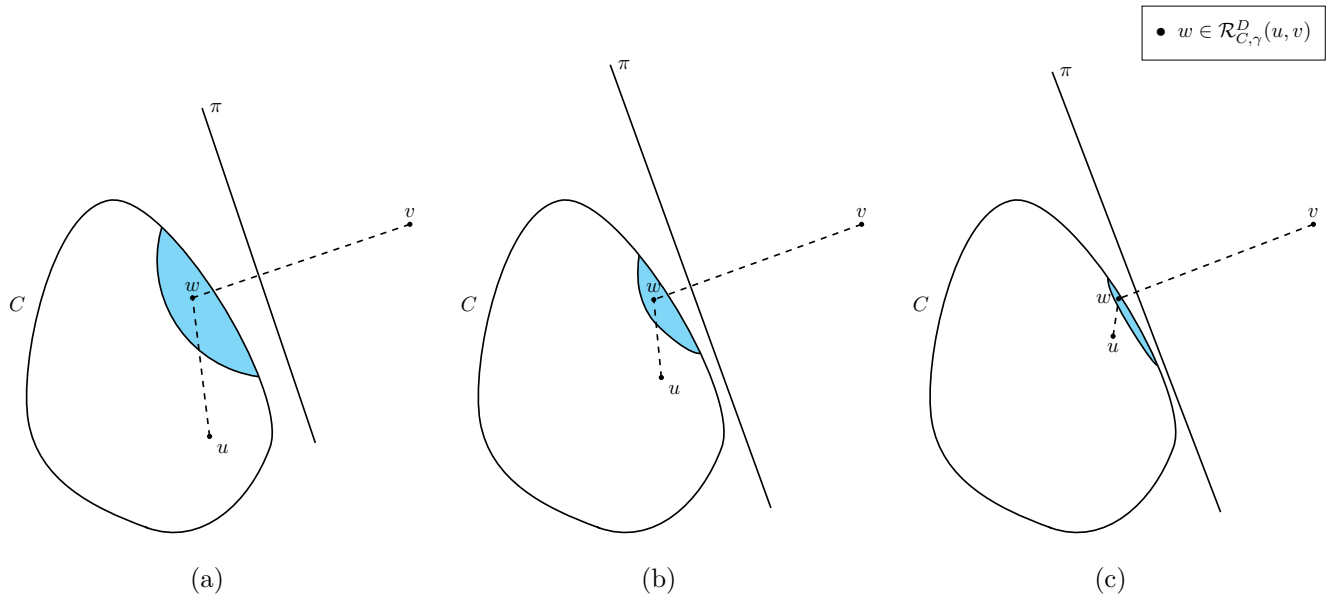


Figure 2.4: Examples of regions given by inexact projection $\mathcal{R}^D_{C,\gamma}(u,v)$.

The next example shows the behavior of the projection $\mathcal{R}^D_{C,\gamma}(u,v)$ when the set C has a vertice and some flat parts.

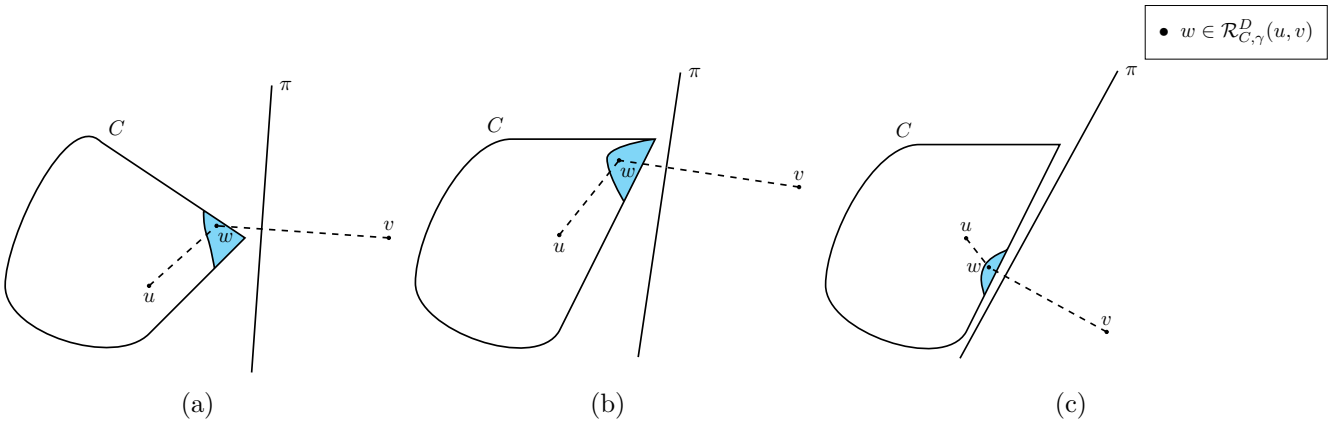


Figure 2.5: Examples of regions given by inexact projection $\mathcal{R}^D_{C,\gamma}(u,v)$.

The next lemma is a variation of [32, Lemma 6]. It will allow to relate Definitions 2.5 and 2.10.

Lemma 2.13. *Let $v \in \mathbb{R}^n$, $u \in C$, $\gamma \geq 0$ and $w \in \mathcal{R}^D_{C,\gamma}(u,v)$. Then, for all $0 \leq \gamma < 1/2$, the holds*

$$\|w-x\|_D^2 \leq \|x-v\|_D^2 + \frac{2\gamma}{1-2\gamma}\|u-v\|_D^2 - \frac{1}{1-2\gamma}\|w-v\|_D^2, \qquad \forall x \in C.$$

Proof. First note that $\|w-x\|_D^2 = \|x-v\|_D^2 - \|w-v\|_D^2 + 2\langle D(v-w), x-w\rangle$. Since $w \in \mathcal{R}^D_{C,\gamma}(u,v)$ and $x \in C$, combining the last equality with (2.5), we obtain

$$\|w-x\|_D^2 \leq \|x-v\|_D^2 - \|w-v\|_D^2 + 2\gamma\|w-u\|_D^2. \tag{2.7}$$?{eq:fg}?

On the other hand, we also have $\|w-u\|_D^2 = \|u-v\|_D^2 - \|w-v\|_D^2 + 2\langle D(v-w), u-w\rangle$. Due to $w \in \mathcal{R}_{C,\gamma}^D(u,v)$ and $u \in C$, using (2.5) and considering that $0 \leq \gamma < 1/2$, we have

$$\|w-u\|_D^2 \leq \frac{1}{1-2\gamma}\|u-v\|_D^2 - \frac{1}{1-2\gamma}\|w-v\|_D^2.$$

Therefore, substituting the last inequality into (2.7), we obtain the desired inequality. □

2.3 The relationship among feasible inexact projections

In the following lemma, we present a relationship between Definitions 2.5 and 2.10.

Lemma 2.14. *Let $v \in \mathbb{R}^n$, $u \in C$, $\gamma \geq 0$ and $\zeta \in (0,1]$. If $0 \leq \gamma < 1/2$ and $\zeta = 1-2\gamma$, then*

$$\mathcal{R}_{C,\gamma}^D(u,v) \subset \mathcal{P}_{C,\zeta}^D(u,v).$$

Proof. Let $w \in \mathcal{R}_{C,\gamma}^D(u,v)$. Applying Lemma 2.13 with $x = \mathcal{P}_C^D(v)$ we have

$$\|w-\mathcal{P}_C^D(v)\|_D^2 \leq \|v-\mathcal{P}_C^D(v)\|_D^2 + \frac{2\gamma}{1-2\gamma}\|u-v\|_D^2 - \frac{1}{1-2\gamma}\|w-v\|_D^2,$$

After some algebraic manipulations in the last inequality we obtain that

$$\|w-v\|_D^2 \leq (1-2\gamma)\|v-\mathcal{P}_C^D(v)\|_D^2 + 2\gamma\|u-v\|_D^2 - (1-2\gamma)\|w-\mathcal{P}_C^D(v)\|_D^2.$$

Therefore, considering that $0 \leq \gamma < 1/2$ and $\zeta = 1-2\gamma$, the result follows from Definition 2.5. □

Remark 2.15. Under the conditions of Lemma 2.14, there exists $0 \leq \gamma < 1/2$ and $\zeta = 1-2\gamma$ such that $\mathcal{P}_{C,\zeta}^D(u,v) \nsubseteq \mathcal{R}_{C,\gamma}^D(u,v)$. Indeed, let $\gamma = 3/8$, $\zeta = 1/4$, and $\bar{w} = \frac{1}{2}(\mathcal{P}_C^D(v)+u)$, then

$$\|\bar{w}-v\|_D^2 = \frac{1}{4}\|\mathcal{P}_C^D(v)-v\|_D^2 + \frac{1}{4}\|u-v\|_D^2 + \frac{1}{2}\langle D(\mathcal{P}_C^D(v)-v), u-v\rangle.$$

Since $\mathcal{P}_C^D(v)$ is the exact projection of v, we have $\langle D(\mathcal{P}_C^D(v)-v), u-v\rangle \leq \|u-v\|_D^2$. Combining this inequality with the last equality and Definition 2.5, we conclude that $\bar{w} \in \mathcal{P}_{C,\zeta}^D(u,v)$. Now, letting $w_t = t\mathcal{P}_C^D(v) + (1-t)\bar{w}$ with $0 < t < 1$, after some algebraic manipulations we have

$$\langle D(v-\bar{w}), w_t - \bar{w}\rangle = t\|\bar{w}-u\|_D^2 - \frac{t}{2}\langle D(v-\mathcal{P}_C^D(v)), u-\mathcal{P}_C^D(v)\rangle.$$

Thus, it follows from Lemma 2.2 that $\langle D(v-\bar{w}), w_t-\bar{w}\rangle \geq t\|\bar{w}-u\|_D^2$. Hence, taking $t > 3/8$ we conclude that $\bar{w} \notin \mathcal{R}_{C,\gamma}^D(u,v)$. Therefore, considering that $\bar{w} \in \mathcal{P}_{C,\zeta}^D(u,v)$, the statement follows.

It follows from Remark 2.15 that, in general, $\mathcal{P}_{C,\zeta}^D(u,v) \nsubseteq \mathcal{R}_{C,\gamma}^D(u,v)$. However, whenever C is a bounded set, we will show that for each fixed $0 \leq \gamma < 1/2$ there exist $0 < \zeta < 1$ such that $\mathcal{P}_{C,\zeta}^D(u,v) \subseteq \mathcal{R}_{C,\gamma}^D(u,v)$. For that, we first need the next lemma.

Lemma 2.16. *Let $v \in \mathbb{R}^n$, $u \in C$ and $0 < \gamma < 1/2$. Assume that C is a bounded set and take*

$$0 < \varepsilon < \frac{\gamma \|u - \mathcal{P}_C^D(v)\|_D^2}{1 - \gamma + \|v - \mathcal{P}_C^D(v)\|_D + 2\gamma\|u - \mathcal{P}_C^D(v)\|_D + \text{diam}C}, \tag{2.8}$$

where diam$C$ *denotes the diameter of C. Then, $\{w \in C : \ \|w - \mathcal{P}_C^D(v)\|_D \leq \varepsilon\} \subset \mathcal{R}_{C,\gamma}^D(u, v)\}$.*

Proof. Take ε satisfying (2.8) and $w \in C$ such that $\|w - \mathcal{P}_C^D(v)\|_D \leq \varepsilon$. For all $z \in C$, we have

$$\begin{aligned}\langle D(v - w), z - w\rangle = \langle D(v - \mathcal{P}_C^D(v)), z - \mathcal{P}_C^D(v)\rangle + \langle D(v - \mathcal{P}_C^D(v)), \mathcal{P}_C^D(v) - w\rangle \\ + \langle D(\mathcal{P}_C^D(v) - w), z - \mathcal{P}_C^D(v)\rangle + \|\mathcal{P}_C^D(v) - w\|_D^2.\end{aligned}$$

Using Lemma 2.2, we have $\langle D(v - \mathcal{P}_C^D(v)), z - \mathcal{P}_C^D(v)\rangle \leq 0$. Thus, the last equality becomes

$$\langle D(v - w), z - w\rangle \leq \langle D(v - \mathcal{P}_C^D(v)), \mathcal{P}_C^D(v) - w\rangle + \langle D(\mathcal{P}_C^D(v) - w), z - \mathcal{P}_C^D(v)\rangle + \|\mathcal{P}_C^D(v) - w\|_D^2.$$

By using Cauchy-Schwarz inequality, we conclude from the last inequality that

$$\langle D(v - w), z - w\rangle \leq \|w - \mathcal{P}_C^D(v)\|_D \left(\|v - \mathcal{P}_C^D(v)\|_D + \|z - \mathcal{P}_C^D(v)\|_D\right) + \|w - \mathcal{P}_C^D(v)\|_D^2.$$

Since $\|w - \mathcal{P}_C^D(v)|_D \leq \varepsilon$ and $\|z - \mathcal{P}_C^D(v)\|_D \leq \text{diam}C$, the last inequality implies that

$$\langle D(v - w), z - w\rangle \leq \varepsilon \left(\|v - \mathcal{P}_C^D(v)\|_D + \text{diam}C\right) + \varepsilon^2, \tag{2.9}$$

On the other hand, if ε satisfies (2.8) then

$$\varepsilon \left(1 - \gamma + \|v - \mathcal{P}_C^D(v)\|_D + \text{diam}C\right) + \gamma\varepsilon^2 < \gamma\|u - \mathcal{P}_C^D(v)\|_D^2 - 2\gamma\varepsilon\|u - \mathcal{P}_C^D(v)\|_D + \gamma\varepsilon^2,$$

hence $\varepsilon \left(1 - \gamma + \|v - \mathcal{P}_C^D(v)\|_D + \text{diam}C\right) + \gamma\varepsilon^2 < \gamma \left(\|u - \mathcal{P}_C^D(v)\|_D - \varepsilon\right)^2$. Since $\gamma, \varepsilon < 1$, we have $\varepsilon^2 < \varepsilon(1 - \gamma) + \gamma\varepsilon^2$ and we can conclude that

$$\varepsilon \left(\|v - \mathcal{P}_C^D(v)\|_D + \text{diam}C\right) + \varepsilon^2 < \gamma \left(\|u - \mathcal{P}_C^D(v)\|_D - \varepsilon\right)^2.$$

It follows from (2.9) that

$$\langle D(v - w), z - w\rangle \leq \gamma \left(\|u - \mathcal{P}_C^D(v)\|_D - \varepsilon\right)^2. \tag{2.10}$$

Using again that $\|w - \mathcal{P}_C^D(v)|_D \leq \varepsilon$ and the triangular inequality, we have

$$0 < \|u - \mathcal{P}_C^D(v)\|_D - \varepsilon \leq \|u - \mathcal{P}_C^D(v)\|_D - \|w - \mathcal{P}_C^D(v)\|_D \leq \|u - w\|_D.$$

Hence, taking into account (2.10), we conclude that $\langle D(v - w), z - w\rangle \leq \gamma\|u - w\|_D^2$. Therefore, it follows from Definition 2.10 that $w \in \mathcal{R}_{C,\gamma}^D(u, v)$. □

Proposition 2.17. *Let $v \in \mathbb{R}^n$, $u \in C$ and assume that C is a bounded set. Then, for each $0 < \gamma < 1/2$, there exist $0 < \zeta < 1$ such that $\mathcal{P}_{C,\zeta}^D(u, v) \subseteq \mathcal{R}_{C,\gamma}^D(u, v)$.*

Proof. It follows from Lemma 2.16 that given $0 < \gamma < 1/2$ there exists $\varepsilon > 0$ such that, for all $w \in C$ with $\|w - \mathcal{P}_C^D(v)\| \leq \varepsilon$, we have $w \in \mathcal{R}_\gamma^D(v)$. Otherwise, we can see in (2.3), when $\zeta \to 1$, the diameter of $C \cap \mathcal{P}_{C,\zeta}^D(u,v)$ tends to zero, then there exists ζ close to 1 such that $\operatorname{diam}(C \cap \mathcal{P}_{C,\zeta}^D(u,v)) < \varepsilon/2$, and $\mathcal{P}_{C,\zeta}^D(u,v) \subset \mathcal{R}_{C,\gamma}^D(u,v)$. □

Next, we present important properties of inexact projections, it will be useful in the sequel.

Lemma 2.18. *Let $x \in C$, $\alpha > 0$ and $z(\alpha) = x - \alpha D^{-1}\nabla f(x)$. Take $w(\alpha) \in \mathcal{P}_{C,\zeta}^D(x, z(\alpha))$ with $\zeta \in (0,1]$. Then,*

(i) $\langle \nabla f(x), w(\alpha) - x\rangle \leq -\dfrac{1}{2\alpha}\|w(\alpha) - x\|_D^2 + \dfrac{\zeta}{2\alpha}\left[\|\mathcal{P}_C^D(z(\alpha)) - z(\alpha)\|_D^2 - \|x - z(\alpha)\|_D^2\right]$;

(ii) the point x is stationary for problem (1) if and only if $x \in \mathcal{P}_{C,\zeta}^D(x, z(\alpha))$;

(iii) if $x \in C$ is a nonstationary point for problem (1), then $\left\langle \nabla f(x), w(\alpha) - x \right\rangle < 0$. *Equivalently, if there exists $\bar{\alpha} > 0$ such that* $\left\langle \nabla f(x), w(\bar{\alpha}) - x \right\rangle \geq 0$, *then x is stationary for problem (1).*

Proof. Since $w(\alpha) \in \mathcal{P}_{C,\zeta}^D(x, z(\alpha))$, applying Lemma 2.9 with $w = w(\alpha)$, $v = z(\alpha)$, $y = x$, and $u = x$, we conclude, after some algebraic manipulations, that

$$\langle D(z(\alpha) - w(\alpha)), x - w(\alpha)\rangle \leq \frac{1}{2}\|w(\alpha) - x\|_D^2 + \frac{\zeta}{2}\left[\|\mathcal{P}_C^D(z(\alpha)) - z(\alpha)\|_D^2 - \|x - z(\alpha)\|_D^2\right].$$

Substituting $z(\alpha) = x - \alpha\nabla f(x)$ in the left hand side of the last inequality, some manipulations yield the inequality of item (i). For proving item (ii), we first assume that x is stationary for problem (1). In this case, (1.1) implies that $\langle \nabla f(x), w(\alpha) - x\rangle \geq 0$. Hence, due to $\|\mathcal{P}_C^D(z(\alpha)) - z(\alpha)\|_D \leq \|x - z(\alpha)\|_D$, item (i) implies

$$\frac{1}{2\alpha}\|w(\alpha) - x\|_D^2 \leq \frac{\zeta}{2\alpha}\left[\|\mathcal{P}_C^D(z(\alpha)) - z(\alpha)\|_D^2 - \|x - z(\alpha)\|_D^2\right] \leq 0.$$

Since $\alpha > 0$ and $\zeta \in (0,1]$, the last inequality yields $w(\alpha) = x$. Therefore, $x \in \mathcal{P}_{C,\zeta}^D(x, z(\alpha))$. Reciprocally, if $x \in \mathcal{P}_{C,\zeta}^D(x, z(\alpha))$, then Definition 2.5 implies that

$$\|x - z(\alpha)\|_D^2 \leq \zeta\|\mathcal{P}_C^D(z(\alpha)) - z(\alpha)\|_D^2 + (1 - \zeta)\|x - z(\alpha)\|_D^2.$$

Hence, $0 \leq \zeta\left(\|\mathcal{P}_C^D(z(\alpha)) - z(\alpha)\|_D^2 - (\|x - z(\alpha)\|_D^2\right)$. Considering that $\zeta \in (0,1]$ we have

$$\|x - z(\alpha)\|_D \leq \|\mathcal{P}_C^D(z(\alpha)) - z(\alpha)\|_D.$$

Thus, due to exact projection with respect to the norm $\|\cdot\|_D$ be unique and $z(\alpha) = x - D^{-1}\alpha\nabla f(x)$, we have $\mathcal{P}_C^D(x - \alpha D^{-1}\nabla f(x)) = x$. Hence, x is the solution of the constrained optimization problem $\min_{y \in C}\|y - z(\alpha)\|_D^2$, which taking into account that $\alpha > 0$ implies (1.1). Therefore, x is stationary point for problem (1). Finally, to prove item (iii), take x a nonstationary point for problem (1). Thus, by item (ii), $x \notin \mathcal{P}_{C,\zeta}^D(x, z(\alpha))$ and taking into account that $w(\alpha) \in \mathcal{P}_{C,\zeta}^D(x, z(\alpha))$, we conclude that $x \neq w(\alpha)$. Since $\|\mathcal{P}_C^D(z(\alpha)) - z(\alpha)\|_D \leq \|x - z(\alpha)\|_D$, $\alpha > 0$ and $\zeta \in (0,1]$, it follows from item (i) that $\left\langle \nabla f(x), w(\alpha) - x \right\rangle < 0$ and the first sentence is proved. Finally, note that the second sentence is the contrapositive of the first sentence. □

Finally, it is worth mentioning that Definitions 2.5 and 2.10, introduced respectively in [15] and [30], are relative inexact concepts, while the concept introduced in [66, 71] is absolute.

2.4 Practical computation of inexact projections

In this section, for a given $v \in \mathbb{R}^n$ and $u \in C$, we discuss how to calculate a point $w \in C$ belonging to $\mathcal{P}^D_{C,\zeta}(u,v)$ or $\mathcal{R}^D_{C,\gamma}(u,v)$. We recall that Lemma 2.14 implies that $\mathcal{P}^D_{C,\zeta}(u,v)$ has more latitude than $\mathcal{R}^D_{C,\gamma}(u,v)$, i.e., $\mathcal{R}^D_{C,\gamma}(u,v) \subset \mathcal{P}^D_{C,\zeta}(u,v)$.

We begin our discussion by showing how a point $w \in \mathcal{P}^D_{C,\zeta}(u,v)$ can be calculated without knowing the point $\mathcal{P}^D_C(v)$. Considering that this discussion has already been covered in [15, Section 3, Algorithm 3.1], we will limit ourselves to giving a general idea of how this task is carried out; see also [18, Section 5.1]. The idea is to use an external procedure capable of computing two sequences $(c_\ell)_{\ell\in\mathbb{N}} \subset \mathbb{R}$ and $(w^\ell)_{\ell\in\mathbb{N}} \subset C$ satisfying the following conditions

$$c_\ell \leq \|\mathcal{P}^D_C(v) - v\|^2_D, \quad \forall \ell \in \mathbb{N}, \qquad \lim_{\ell\to+\infty} c_\ell = \|\mathcal{P}^D_C(v) - v\|^2_D, \qquad \lim_{\ell\to+\infty} w^\ell = \mathcal{P}^D_C(v). \tag{2.11}$$?{def:cl}

In this case, if $v \notin C$, then we have $\|\mathcal{P}^D_C(v) - v\|^2_D - \|u - v\|^2_D < 0$. Hence, given an arbitrary $\zeta \in (0,1)$, the second condition in (2.11) implies that there exists $\hat{\ell}$ such that

$$\|\mathcal{P}^D_C(v) - v\|^2_D - \|u - v\|^2_D < \zeta(c_{\hat{\ell}} - \|u - v\|^2_D).$$

Moreover, by using the last condition in (2.11), we conclude that there exists $\bar{\ell} > \hat{\ell}$ such that

$$\|w_{\bar{\ell}} - v\|^2_D - \|u - v\|^2_D < \zeta(c_{\bar{\ell}} - \|u - v\|^2_D), \tag{2.12}$$?{def:cls}

which using the inequality in (2.11) yields $\|w_{\bar{\ell}} - v\|^2_D < \zeta\|\mathcal{P}^D_C(v) - v\|^2_D + (1-\zeta)\|u - v\|^2_D$. Hence, Definition 2.5 implies that $w_{\bar{\ell}} \in \mathcal{P}^D_{C,\zeta}(u,v)$. Therefore, (2.12) can be used as a stopping criterion to compute a feasible inexact projection, with respect to the norm $\|\cdot\|_D$, of v onto C relative to u and forcing parameter $\zeta \in (0,1]$. For instance, it follows from [15, Theorem 3.2, Lemma 3.1] (see also [17]) that such sequences $(c_\ell)_{\ell\in\mathbb{N}} \subset \mathbb{R}$ and $(w^\ell)_{\ell\in\mathbb{N}} \subset C$ satisfying (2.11) can be computed by using *Dykstra's algorithm* [24, 34], whenever D is the identity matrix and the set $C = \cap_{i=1}^p C_i$, where C_i are closed and convex sets and the exact projection $\mathcal{P}^D_{C_i}(v)$ is easy to obtain, for all $i = 1, \ldots, p$.

We end this section by discussing how to compute a point $w \in \mathcal{R}^D_{C,\gamma}(u,v)$. For that, we apply the classical *Frank-Wolfe method*, also known as *conditional gradient method*, to minimize the function $\psi(z) := \|z - v\|^2/2$ onto the constraint set C with a suitable stop criteria depending of $u \in C$ and $\gamma \in (0,1]$, see [10, 51]. To state the method we assume the existence of a linear optimization oracle (or simply LO oracle) capable of minimizing linear functions over the constraint set C, which is assumed to be compact. The Frank-Wolfe method is formally stated as follows.

Algorithm 2.1 Frank-Wolfe method to compute $w \in \mathcal{R}^D_{C,\gamma}(u,v)$

Input: $D \in \mathcal{D}_\mu$, $\gamma \in (0,1]$, $v \in \mathbb{R}^n$ and $u \in C$.

Step 0. Let $w^0 \in C$ and set $\ell \leftarrow 0$.

Step 1. Use a LO oracle to compute an optimal solution z^ℓ and the optimal value s^*_ℓ as

$$z^\ell \in \arg\min_{z\in C} \langle w^\ell - v,\ z - w^\ell\rangle, \qquad s^*_\ell := \langle w^\ell - v,\ z^\ell - w^\ell\rangle. \tag{2.13}$$?{eq:Cond

If $-s^*_\ell \le \gamma\|w^\ell - u\|^2_D$, then define $w := w^\ell$ and **stop**.

Step 2. Compute α_ℓ and $w_{\ell+1}$ as

$$w_{\ell+1} := w^\ell + \alpha_\ell(z^\ell - w^\ell), \qquad \alpha_\ell := \min\left\{1, -s^*_\ell/\|z^\ell - w^\ell\|^2\right\}. \tag{2.14}$$?{eq:step

Set $\ell \leftarrow \ell + 1$, and go to Step 1.

Output: $w := w^\ell$.

Let us describe the main features of Algorithm 2.1, i.e., the Frank-Wolfe method applied to the problem $\min_{z\in C}\psi(z)$. In this case, (2.13) is equivalent to $s^*_\ell := \min_{z\in C}\langle\psi'(w^\ell),\ z - w^\ell\rangle$. Since ψ is convex, we have $\psi(z) \ge \psi(w^\ell) + \langle\psi'(w^\ell),\ z - w^\ell\rangle \ge \psi(w^\ell) + s^*_\ell$, for all $z \in C$. Define $w_* := \arg\min_{z\in C}\psi(z)$ and $\psi^* := \min_{z\in C}\psi(z)$. Letting $z = w_*$ in the last inequality, we obtain $\psi(w^\ell) \ge \psi^* \ge \psi(w^\ell) + s^*_\ell$, which implies that $s^*_\ell < 0$ whenever $\psi(w^\ell) \ne \psi^*$. Thus, we conclude that $-s^*_\ell = \langle v - w^\ell,\ z^\ell - w^\ell\rangle > 0 \ge \langle v - w_*,\ z - w_*\rangle$, for all $z \in C$. Therefore, if Algorithm 2.1 computes $w^\ell \in C$ satisfying $-s^*_\ell \le \gamma\|w^\ell - u\|^2_D$, then the method terminates. Otherwise, it computes the step size $\alpha_\ell = \arg\min_{\alpha\in[0,1]}\psi(w^\ell + \alpha(z^\ell - w^\ell))$ using exact minimization. Since z^ℓ, $w^\ell \in C$ and C is convex, we conclude from (2.14) that $w_{\ell+1} \in C$, thus Algorithm 2.1 generates a sequence in C. Finally, (2.13) implies that $\langle v - w^\ell,\ z - w^\ell\rangle \le -s^*_\ell$, for all $z \in C$. Considering that [10, Proposition A.2] implies that $\lim_{\ell\to+\infty} s^*_\ell = 0$ and taking into account the stopping criteria $-s^*_\ell \le \gamma\|w^\ell - u\|^2_D$, we conclude that the output of Algorithm 2.1 is a feasible inexact projection $w \in \mathcal{R}^D_{C,\gamma}(u,v)$ i.e., $\langle v - w,\ z - w\rangle \le \gamma\|w^\ell - u\|^2_D$, for all $z \in C$.

Chapter 3

Inexact scaled gradient method

The aim of this chapter is to present an inexact version of SGP method, which inexactness are in two distinct senses. First, we use a version of the inexactness scheme introduced in [15], and also a variation of the one appeared in [71], to compute an inexact projection onto the feasible set allowing an appropriate relative error tolerance. Second, using the inexactness conceptual scheme for nonmonotones line search introduced in [46, 65], a step size is computed to define the next iterate. The statement of the conceptual algorithm is as follows.

Algorithm 3.1 SGP with inexact projection and nonmonotone line search

Step 0. Choose $\sigma, \zeta_{\min} \in (0,1)$, $\delta_{\min} \in [0,1)$, $0 < \underline{\omega} < \bar{\omega} < 1$, $0 < \alpha_{\min} \leq \alpha_{\max}$ and $\mu \geq 1$. Let $x^0 \in C$, $\nu_0 \geq 0$ and set $k \leftarrow 0$.

Step 1. Choose positive real numbers α_k and ζ_k, and a positive definite matrix D_k such that

$$\alpha_{\min} \leq \alpha_k \leq \alpha_{\max}, \qquad 0 < \zeta_{\min} < \zeta_k \leq 1, \qquad D_k \in \mathcal{D}_\mu. \tag{3.1}$$?{eq:TolA

Compute $w^k \in C$ as any feasible inexact projection with respect to the norm $\|\cdot\|_{D_k}$ of $z^k := x^k - \alpha_k D_k^{-1} \nabla f(x^k)$ onto C relative to x^k with forcing parameter ζ_k, i.e.,

$$w^k \in \mathcal{P}_{C,\zeta_k}^{D_k}(x^k, z^k). \tag{3.2}$$?{eq:PIne

If $w^k = x^k$, then **stop** declaring convergence.

Step 2. Set $\tau_{\text{trial}} \leftarrow 1$. If

$$f\Big(x^k + \tau_{\text{trial}}(w^k - x^k)\Big) \leq f(x^k) + \sigma \tau_{\text{trial}} \Big\langle \nabla f(x^k), w^k - x^k \Big\rangle + \nu_k, \tag{3.3}$$?{eq:TkAr

then $\tau_k \leftarrow \tau_{\text{trial}}$, define the next iterate x^{k+1} as

$$x^{k+1} = x^k + \tau_k (w^k - x^k), \tag{3.4}$$?{eq:Iter

and go to **Step 3**. Otherwise, choose $\tau_{\text{new}} \in [\underline{\omega}\tau_{\text{trial}}, \bar{\omega}\tau_{\text{trial}}]$, set $\tau_{\text{trial}} \leftarrow \tau_{\text{new}}$, and repeat test (3.3).

Step 3. Take $\delta_{k+1} \in [\delta_{\min}, 1]$ and choose $\nu_{k+1} \in \mathbb{R}$ satisfying

$$0 \leq \nu_{k+1} \leq (1 - \delta_{k+1})\Big[f(x^k) + \nu_k - f(x^{k+1})\Big]. \tag{3.5}$$?{eq:nuk}

Set $k \leftarrow k+1$ and go to **Step 1**.

Let us describe the main features of Algorithm 3.1. In Step 1, we first choose $\alpha_{\min} \leq \alpha_k \leq \alpha_{\max}$, $0 < \zeta_{\min} \leq \zeta_k < 1$, and $D_k \in \mathcal{D}_\mu$. Then, by using some (inner) procedure, such as those specified in Section 2, we compute w^k as any feasible inexact projection of $z^k = x_k - \alpha_k D_k^{-1} \nabla f(x_k)$ onto the feasible set C relative to the previous iterate x^k with forcing parameter ζ_k. If $w^k = x^k$, then Lemma 2.18*(ii)* implies that x^k is a solution of problem (1). Otherwise, $w^k \neq x^k$ and Lemma 2.18*(i)* implies that $w^k - x^k$ is a descent direction of f at x^k, i.e., $\langle \nabla f(x^k), w^k - x^k \rangle < 0$. Hence, in Step 2, we employ a nonmonotone line search with tolerance parameter $\nu_k \geq 0$ to compute a step size $\tau_k \in (0,1]$, and the next iterate is computed as in (3.4). Finally, due to (3.3) and $\delta_{k+1} \in [\delta_{\min}, 1]$, we have $0 \leq (1 - \delta_{k+1})\Big[f(x^k) + \nu_k - f(x^{k+1})\Big]$. Therefore, the next tolerance parameter $\nu_{k+1} \in \mathbb{R}$ can be chosen satisfying (3.5) in Step 3, completing the iteration.

It is worth mentioning that the conditions in (3.1) allow combining several strategies for choosing the step sizes α_k and the matrices D_k to accelerate the performance of the classical gradient method. Strategies of choosing the step sizes α_k and the matrices D_k have their origin in the study of the gradient method for unconstrained optimization, papers dealing with this issue include but

are not limited to [8, 29, 31, 38, 76], see also [20, 27, 28, 52]. More details about selecting step sizes α_k and matrices D_k can be found in the recent review [19] and references therein.

Below, we present some particular instances of the parameter $\delta_k \geq 0$ and the non-monotonicity tolerance parameter $\nu_k \geq 0$ in Step 3.

1. *Armijo line search*

   Taking $\nu_k \equiv 0$, the line search (3.3) is the well-known (monotone) Armijo line search, see [13, Section 2.3]. In this case, we can take $\delta_k \equiv 1$ in Step 3.

2. *Max-type line search*

   The earliest nonmonotone line search strategy was proposed in [48]. Let $M > 0$ be an integer parameter. In an iteration k, this strategy requires a step size $\tau_k > 0$ satisfying

   $$f\Big(x^k + \tau_k(w^k - x^k)\Big) \leq \max_{0\leq j\leq m_k} f(x^{k-j}) + \sigma\tau_k\Big\langle \nabla f(x^k), w^k - x^k\Big\rangle, \tag{3.6}$$?{eq:grip

   where $m_0 = 0$ and $0 \leq m_k \leq \min\{m_{k-1} + 1, M\}$. To simplify the notations, we define $f(x^{\ell(k)}) := \max_{0\leq j\leq m_k} f(x^{k-j})$. In order to identify (3.6) as a particular instance of (3.3), we set

   $$\nu_k = f(x^{\ell(k)}) - f(x^k), \quad 0 = \delta_{\min} \leq \delta_{k+1} \leq [f(x^{\ell(k)}) - f(x^{\ell(k+1)})]/[f(x^{\ell(k)}) - f(x^{k+1})]. \tag{3.7}$$?{eq:casg

   Parameters ν_k and δ_{k+1} in (3.7) satisfy the corresponding conditions in Algorithm 3.1, i.e., $\nu_k \geq 0$ and $\delta_{k+1} \in [\delta_{\min}, 1]$ (with $\delta_{\min} = 0$) satisfy (3.5). In fact, the definition of $f(x^{\ell(k)})$ implies that $f(x^k) \leq f(x^{\ell(k)})$ and hence $\nu_k \geq 0$. Due to $\langle \nabla f(x^k), w^k - x^k\rangle < 0$, it follows from (3.3) that $f(x^{\ell(k)}) - f(x^{k+1}) > 0$. Since $m_{k+1} \leq m_k + 1$, we conclude that $f(x^{\ell(k)}) - f(x^{\ell(k+1)}) \geq 0$. Hence, since $f(x^{k+1}) \leq f(x^{\ell(k+1)})$, we obtain $\delta_{k+1} \in [0, 1]$. Moreover, (3.5) is equivalent to

   $$\delta_{k+1}[f(x^k) + \nu_k - f(x^{k+1})] \leq (f(x^k) + \nu_k) - (f(x^{k+1}) + \nu_{k+1}),$$

   which in turn, taking into account that $\nu_k = f(x^{\ell(k)}) - f(x^k)$, is equivalent to second inequality in (3.7). Thus, (3.6) is a particular instance of (3.3) with ν_k and δ_{k+1} defined in (3.7). Therefore, Algorithm 3.1 has as a particular instance the inexact projected version of the scaled gradient method employing the nonmonotone line search (3.6). This version has been considered in [15]; see also [21, 72].

3. *Average-type line search*

   Let us first recall the definition of the sequence of "cost updates" $(c_k)_{k\in\mathbb{N}}$ that characterizes the nonmonotone line search proposed in [75]. Let $0 \leq \eta_{\min} \leq \eta_{\max} < 1$, $c_0 = f(x_0)$ and $q_0 = 1$. Choose $\eta_k \in [\eta_{\min}, \eta_{\max}]$ and set

   $$q_{k+1} = \eta_k q_k + 1, \qquad c_{k+1} = [\eta_k q_k c_k + f(x^{k+1})]/q_{k+1}, \qquad \forall k \in \mathbb{N}. \tag{3.8}$$?{eq:zhs}

Some algebraic manipulations show that the sequence defined in (3.8) is equivalent to

$$c_{k+1} = (1 - 1/q_{k+1})c_k + f(x^{k+1})/q_{k+1}, \qquad \forall k \in \mathbb{N}. \tag{3.9}$$?{eq:zhsn

Since (3.5) is equivalent to

$$f(x^{k+1}) + \nu_{k+1} \leq (1 - \delta_{k+1})(f(x^k) + \nu_k) + \delta_{k+1} f(x^{k+1}),$$

it follows from (3.9) that letting $\nu_k = c_k - f(x^k)$ and $\delta_{k+1} = 1/q_{k+1}$, Algorithm 3.1 becomes the inexact projected version of the scaled gradient method employing the nonmonotone line search proposed in [75]. Finally, considering that $q_0 = 1$ and $\eta_{\max} < 1$, the first equality in (3.8) implies that

$$q_{k+1} = 1 + \sum_{j=0}^{k} \prod_{i=0}^{j} \eta_{k-i} \leq \sum_{j=0}^{+\infty} \eta_{\max}^j = 1/(1 - \eta_{\max}).$$

In this case, due to $\delta_{k+1} = 1/q_{k+1}$, we can take $\delta_{\min} = 1 - \eta_{\max} > 0$ in Step 3. For gradient projection methods employing the nonmonotone Average-type line search see, for example, [7, 36, 73].

Remark 3.1. The general line search in Step 2 of Algorithm 3.1 with parameters δ_{k+1} and ν_k properly chosen in Step 3, also contains as particular cases the nonmonotone line searches that appeared in [3, 55], see also [46].

3.1 Partial asymptotic convergence analysis

The goal of this section is to present a partial convergence result for the sequence $(x^k)_{k\in\mathbb{N}}$ generated by Algorithm 3.1, namely, we will prove that every cluster point of $(x^k)_{k\in\mathbb{N}}$ is stationary for problem (1). For that, we state a result that is contained in the proof of [46, Theorem 4].

Lemma 3.2. *For all $k \in \mathbb{N}$,*

$$0 \leq \delta_{k+1}\Big[f(x^k) + \nu_k - f(x^{k+1})\Big] \leq \Big(f(x^k) + \nu_k\Big) - \Big(f(x^{k+1}) + \nu_{k+1}\Big).$$

As consequence the sequence $\Big(f(x^k) + \nu_k\Big)_{k\in\mathbb{N}}$ is non-increasing.

Next, we present our first convergence result. It is worth noting that, just as in the classical projected gradient method, we do not need to assume that f has a bounded sub-level set.

Proposition 3.3. *Assume that $\lim_{k\to+\infty} \nu_k = 0$. Then, Algorithm 3.1 stops in a finite number of iterations at a stationary point of problem (1), or generates an infinite sequence $(x^k)_{k\in\mathbb{N}}$ for which every cluster point is stationary for problem (1).*

Proof. First, assume that $(x^k)_{k\in\mathbb{N}}$ is finite. In this case, according to Step 1, there exists $k \in \mathbb{N}$ such that $x^k = w^k \in \mathcal{P}^{D_k}_{C,\zeta_k}(x^k, z^k)$, where $z^k = x^k - \alpha_k D_k^{-1}\nabla f(x^k)$, $0 < \bar{\zeta} < \zeta_k \leq 1$ and $\alpha_k > 0$.

Therefore, applying Lemma 2.18*(ii)* with $x = x^k$, $\alpha = \alpha_k$ and $\zeta = \zeta_k$, we conclude that x^k is stationary for problem (1). Now, assume that $(x^k)_{k\in\mathbb{N}}$ is infinite. Let $\bar{x}$ be a cluster point of $(x^k)_{k\in\mathbb{N}}$ and $(x^{k_j})_{j\in\mathbb{N}}$ be a subsequence of $(x^k)_{k\in\mathbb{N}}$ such that $\lim_{j\to+\infty} x^{k_j} = \bar{x}$. Since C is closed and $(x^k)_{k\in\mathbb{N}} \subset C$, we have $\bar{x} \in C$. Moreover, since $\lim_{k\to+\infty} \nu_k = 0$, we have $\lim_{j\to+\infty} \Big(f(x^{k_j}) + \nu_{k_j}\Big) = f(\bar{x})$. Hence, considering that $\lim_{k\to+\infty} \nu_k = 0$ and Lemma 3.2 implies that $\Big(f(x^k) + \nu_k\Big)_{k\in\mathbb{N}}$ is non-increasing, we conclude that $\lim_{k\to+\infty} f(x^k) = \lim_{k\to+\infty} \Big(f(x^k) + \nu_k\Big) = f(\bar{x})$. On the other hand, due to $w^k \in \mathcal{P}^{D_k}_{C,\zeta_k}(x^k, z^k)$, where $z^k = x^k - \alpha_k \nabla f(x^k)$, Definition 2.5 implies

$$\|w^{k_j} - z^{k_j}\|^2_{D_k} \leq \zeta_{k_j} \|\mathcal{P}^{D_k}_C(z^{k_j}) - z^{k_j}\|^2_{D_k} + (1 - \zeta_{k_j})\|x^{k_j} - z^{k_j}\|^2_{D_k}. \tag{3.10}$$?{eq:bsw}

Considering that $(\alpha_k)_{k\in\mathbb{N}}$ and $(\zeta_k)_{k\in\mathbb{N}}$ are bounded, $(D_k)_{k\in\mathbb{N}} \subset \mathcal{D}_\mu$, $(x^{k_j})_{j\in\mathbb{N}}$ converges to $\bar{x}$ and ∇f is continuous, the last inequality together Remark 2.4 and (1.3) imply that $(w^{k_j})_{j\in\mathbb{N}} \subset C$ is also bounded. Thus, we can assume without loss of generality that $\lim_{j\to+\infty} w^{k_j} = \bar{w} \in C$. In addition, taking into account that $x^k \neq w^k$ for all $k = 0, 1, \ldots$, applying Lemma 2.18*(i)* with $x = x^k$, $\alpha = \alpha_k$, $z(\alpha) = z^k$ and $\zeta = \zeta_k$, we obtain that $\langle \nabla f(x^k), w^k - x^k\rangle < 0$, for all $k = 0, 1, \ldots$. Therefore, (3.3) and (3.4) imply that

$$0 < -\sigma\tau_k \Big\langle \nabla f(x^k), w^k - x^k \Big\rangle \leq f(x^k) + \nu_k - f(x^{k+1}), \qquad \forall\ k \in \mathbb{N}. \tag{3.11}$$?{eq:fmot

Now, due $\tau_k \in (0, 1]$, for all $k = 0, 1, \ldots$, we can also assume without loss of generality that $\lim_{j\to+\infty} \tau_{k_j} = \bar{\tau} \in [0, 1]$. Therefore, since $\lim_{k\to+\infty} f(x^k) = f(\bar{x})$ and $\lim_{k\to+\infty} \nu_k = 0$, taking limit in (3.11) along the subsequences $(x^{k_j})_{j\in\mathbb{N}}$, $(w^{k_j})_{j\in\mathbb{N}}$ and $(\tau_{k_j})_{j\in\mathbb{N}}$ yields $\bar{\tau}\Big\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\Big\rangle = 0$. We have two possibilities: $\bar{\tau} > 0$ or $\bar{\tau} = 0$. If $\bar{\tau} > 0$, then $\Big\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\Big\rangle = 0$. Now, we assume that $\bar{\tau} = 0$. In this case, for all j large enough, there exists $0 < \hat{\tau}_{k_j} \leq \min\{1, \tau_{k_j}/\underline{\omega}\}$ such that

$$f\Big(x^{k_j} + \hat{\tau}_{k_j}(w^{k_j} - x^{k_j})\Big) > f(x^{k_j}) + \sigma\hat{\tau}_{k_j}\Big\langle \nabla f(x^{k_j}), w^{k_j} - x^{k_j}\Big\rangle + \nu_{k_j}. \tag{3.12}$$?{eq:ffA1

On the other hand, by the mean value theorem, there exists $\xi_{k_j} \in (0, 1)$ such that

$$\langle \nabla f\Big(x^{k_j} + \xi_{k_j}\hat{\tau}_{k_j}(w^{k_j} - x^{k_j})\Big), \hat{\tau}_{k_j}(w^{k_j} - x^{k_j})\rangle = f\Big(x^{k_j} + \hat{\tau}_{k_j}(w^{k_j} - x^{k_j})\Big) - f(x^{k_j}).$$

Combining this equality with (3.12), and taking into account that $\nu_{k_j} \geq 0$, we have

$$\langle \nabla f\Big(x^{k_j} + \xi_{k_j}\hat{\tau}_{k_j}(w^{k_j} - x^{k_j})\Big), \hat{\tau}_{k_j}(w^{k_j} - x^{k_j})\rangle > \sigma\hat{\tau}_{k_j}\Big\langle \nabla f(x^{k_j}), w^{k_j} - x^{k_j}\Big\rangle,$$

for j large enough. Since $0 < \hat{\tau}_{k_j} \leq \min\{1, \tau_{k_j}/\underline{\omega}\}$, it follows that $\lim_{j\to\infty} \hat{\tau}_{k_j}\|w^{k_j} - x^{k_j}\| = 0$. Then, dividing both sides of the above inequality by $\hat{\tau}_{k_j} > 0$ and taking limits as j goes to $+\infty$, we conclude that $\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\rangle \geq \sigma\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\rangle$. Hence, due to $\sigma \in (0, 1)$, we obtain $\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\rangle \geq 0$. We recall that $\langle \nabla f(x^{k_j}), w^{k_j} - x^{k_j}\rangle < 0$, for all $j = 0, 1, \ldots$, which taking limit as j goes to $+\infty$ yields $\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\rangle \leq 0$. Hence, we also have $\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\rangle = 0$. Therefore, for any of the two possibilities, $\bar{\tau} > 0$ or $\bar{\tau} = 0$, we have $\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\rangle = 0$. On the other hand, since $(\alpha_k)_{k\in\mathbb{N}}$ and $(\zeta_k)_{k\in\mathbb{N}}$ are bounded, we also assume without loss of generality that $\lim_{j\to+\infty} \alpha_{k_j} = \bar{\alpha} \in [\alpha_{\min}, \alpha_{\max}]$ and $\lim_{j\to+\infty} \zeta_{k_j} = \bar{\zeta} \in [\zeta_{\min}, 1]$. Thus, since Remark 2.4 implies that

$$\lim_{j\to+\infty} \mathcal{P}^{D_{k_j}}_C(z^{k_j}) = \mathcal{P}^{\bar{D}}_C(\bar{z}),$$

and considering that $\lim_{j\to+\infty} x^{k_j} = \bar{x} \in C$, $\lim_{j\to+\infty} w^{k_j} = \bar{w} \in C$, $\lim_{j\to+\infty} \tau_{k_j} = \bar{\tau} \in [0,1]$, $\lim_{j\to+\infty} D_{k_j} = \bar{D} \in \mathcal{D}_\mu$, taking limit in (3.10), we conclude that

$$\|\bar{w} - \bar{z}\|_{\bar{D}}^2 \leq \bar{\zeta}\|\mathcal{P}_C^{\bar{D}}(\bar{z}) - \bar{z}\|_{\bar{D}}^2 + (1-\bar{\zeta})\|\bar{x} - \bar{z}\|_{\bar{D}}^2,$$

where $\bar{z} = \bar{x} - \bar{\alpha}\nabla f(\bar{x})$. Hence, Definition 2.5 implies that $\bar{w} \in \mathcal{P}_{C,\bar{\zeta}}^{\bar{D}}(\bar{x}, \bar{z})$, where $\bar{z} = \bar{x} - \bar{\alpha}\nabla f(\bar{x})$. Therefore, due to $\langle \nabla f(\bar{x}), \bar{w} - \bar{x}\rangle = 0$, we can apply second sentence in Lemma 2.18*(iii)* with $x = \bar{x}$, $z(\bar{\alpha}) = \bar{z}$ and $w(\bar{\alpha}) = \bar{w}$, to conclude that $\bar{x}$ is stationary for problem (1). □

The tolerance parameter ν_k that controls the non-monotonicity of the line search must be smaller and smaller as the sequence $(x^k)_{k\in\mathbb{N}}$ tends to a stationary point. Next corollary presents a general condition for this property, its proof can be found in [46, Theorem 4].

Corollary 3.4. *If $\delta_{\min} > 0$, then $\sum_{k=0}^{+\infty} \nu_k < +\infty$. Consequently, $\lim_{k\to+\infty} \nu_k = 0$.*

The Armijo and the nonmonotone Average-type line searches discussed in Section 3 satisfy the assumption of Corollary 3.4, i.e., $\delta_{\min} > 0$. However, for the nonmonotone Max-type line search, we can only guarantee that $\delta_{\min} \geq 0$. Hence, we can not apply Corollary 3.4 to conclude that $\lim_{k\to+\infty} \nu_k = 0$. In the next proposition, we will deal with this case separately.

Proposition 3.5. *Assume that the sequence $(x^k)_{k\in\mathbb{N}}$ is generated by Algorithm 3.1 with the nonmonotone line search (3.6), i.e., $\nu_k = f(x^{\ell(k)}) - f(x^k)$ for all $k \in \mathbb{N}$. In addition, assume that the level set $C_0 := \{x \in C : \ f(x) \leq f(x^0)\}$ is bounded and $\nu_0 = 0$. Then, $\lim_{k\to+\infty} \nu_k = 0$.*

Proof. First of all, note that $w^k \in \mathcal{P}_{C,\zeta_k}^{D_k}(x^k, z^k)$, where $z^k = x^k - \alpha_k D_k^{-1}\nabla f(x^k)$ and $D_k \in \mathcal{D}_\mu$. Thus, applying Lemma 2.18*(i)* with $x = x^k$, $w(\alpha) = w^k$, $z = z^k$ and $\zeta = \zeta_k$, we obtain

$$\|w^k - x^k\|^2 \leq -2\mu\alpha_{\max}\langle \nabla f(x^k), w^k - x^k\rangle, \qquad \forall k \in \mathbb{N}. \tag{3.13}$$?{eq:apna

On the other hand, due to $f(x^{\ell(k)}) = f(x^k) + \nu_k$, Lemma 3.2 implies that $(f(x^{\ell(k)}))_{k\in\mathbb{N}}$ is non-increasing and

$$f(x^{k+1}) \leq f(x^{k+1}) + \nu_{k+1} \leq f(x^k) + \nu_k \leq f(x^0).$$

Hence, we have $(x^k)_{k\in\mathbb{N}} \subset C_0$ and, as a consequence, $(f(x^{\ell(k)}))_{k\in\mathbb{N}}$ converges. Note that $\ell(k)$ is an integer such that

$$k - m_k \leq \ell(k) \leq k. \tag{3.14}$$?{eq:lk}?

Since $x^{\ell(k)} = x^{\ell(k)-1} + \tau_{\ell(k)-1}(w^{\ell(k)-1} - x^{\ell(k)-1})$, (3.6) implies that

$$f\left(x^{\ell(k)}\right) \leq f\left(x^{\ell(\ell(k)-1)}\right) + \sigma\tau_{\ell(k)-1}\left\langle \nabla f(x^{\ell(k)-1}), w^{\ell(k)-1} - x^{\ell(k)-1}\right\rangle,$$

for all $k > M$. In view of $(f(x^{\ell(k)}))_{k\in\mathbb{N}}$ be convergent, $\langle \nabla f(x^k), w^k - x^k\rangle < 0$ for all $k \in \mathbb{N}$, and taking into account that $\tau_k \in (0,1]$, the last inequality together (3.13) implies that

$$\lim_{k\to+\infty} \tau_{\ell(k)-1}\|w^{\ell(k)-1} - x^{\ell(k)-1}\| = 0. \tag{3.15}$$?{eq:apcs

We proceed to prove that $\lim_{k\to+\infty} f(x^k) = \lim_{k\to+\infty} f(x^{\ell(k)})$. For that, set $\hat{\ell}(k) := \ell(k+M+2)$. First, we prove by induction that, for all $j \geq 1$, the following two equalities hold

$$\lim_{k\to+\infty} \tau_{\hat{\ell}(k)-j}\|w^{\hat{\ell}(k)-j} - x^{\hat{\ell}(k)-j}\| = 0, \qquad \lim_{k\to+\infty} f(x^{\hat{\ell}(k)-j}) = \lim_{k\to+\infty} f(x^{\ell(k)}), \tag{3.16}$$?{eq:ind}

where we are considering $k \geq j-1$. Assume that $j = 1$. Since $\{\hat{\ell}(k) : \ k \in \mathbb{N}\} \subset \{\ell(k) : \ k \in \mathbb{N}\}$, the first equality in (3.16) follows from (3.15). Hence, $\lim_{k\to+\infty} \|x^{\hat{\ell}(k)} - x^{\hat{\ell}(k)-1}\| = 0$. Since C_0 is compact and f is uniformly continuous on C_0, we have $\lim_{k\to+\infty} f(x^{\hat{\ell}(k)-1}) = \lim_{k\to+\infty} f(x^{\hat{\ell}(k)})$, which again using that $\{\hat{\ell}(k) : \ k \in \mathbb{N}\} \subset \{\ell(k) : \ k \in \mathbb{N}\}$ implies the second equality in (3.16). Assume that (3.16) holds for j. Again, due to $x^{\hat{\ell}(k)-j} = x^{\hat{\ell}(k)-j-1} + \tau_{\hat{\ell}(k)-j-1}(w^{\hat{\ell}(k)-j-1} - x^{\hat{\ell}(k)-j-1})$, (3.6) implies that

$$f\Big(x^{\hat{\ell}(k)-j}\Big) \leq f\Big(x^{\ell(\hat{\ell}(k)j-(j+1))}\Big) + \sigma\tau_{\hat{\ell}(k)-(j+1)}\Big\langle \nabla f(x^{\hat{\ell}(k)-(j+1)}), w^{\hat{\ell}(k)-(j+1)} - x^{\hat{\ell}(k)-(j+1)}\Big\rangle.$$

Similar argument used to obtain (3.15) yields $\lim_{k\to+\infty} \tau_{\hat{\ell}(k)-(j+1)}\|w^{\hat{\ell}(k)-(j+1)} - x^{\hat{\ell}(k)-(j+1)}\| = 0$. Thus, the first equality in (3.16) holds for $j+1$, which implies $\lim_{k\to+\infty} \|x^{\hat{\ell}(k)-j} - x^{\hat{\ell}(k)-(1+j)}\| = 0$. Again, the uniformly continuity of f on C_0 gives

$$\lim_{k\to+\infty} f(x^{\hat{\ell}(k)-(j+1)}) = \lim_{k\to+\infty} f(x^{\hat{\ell}(k)-j}),$$

which shows that the second equality in (3.16) holds for $j+1$. From (3.14) and $\hat{\ell}(k) := \ell(k+M+2)$, we obtain $\hat{\ell}(k) - k - 1 \leq M + 1$. Thus, taking into account that

$$x^{k+1} = x^{\hat{\ell}(k)} - \sum_{j=1}^{\hat{\ell}(k)-k-1} \tau_{\hat{\ell}(k)-j}\Big(w^{\hat{\ell}(k)-j} - x^{\hat{\ell}(k)-j}\Big),$$

it follows from the first inequality in (3.16) that $\lim_{k\to+\infty} \|x^{k+1} - x^{\hat{\ell}(k)}\| = 0$. Hence, due to f be uniformly continuous on C_0 and $(f(x^{\ell(k)}))_{k\in\mathbb{N}}$ be convergent, we conclude that

$$\lim_{k\to+\infty} f(x^k) = \lim_{k\to+\infty} f(x^{\hat{\ell}(k)}) = \lim_{k\to+\infty} f(x^{\ell(k)}),$$

and considering that $\nu_k = f(x^{\ell(k)}) - f(x^k)$ the desired results follows. □

Remark 3.6. Let $C_0 := \{x \in C : \ f(x) \leq f(x^0)\}$ be bounded and $(x^k)_{k\in\mathbb{N}}$ be generated by Algorithm 3.1 with the nonmonotone line search (3.6) with $\nu_0 = 0$. Then, combining Propositions 3.3 and 3.5, we conclude that $(x^k)_{k\in\mathbb{N}}$ is either finite terminating at a stationary point of problem (1), or infinite, and every cluster point of $(x^k)_{k\in\mathbb{N}}$ is stationary for problem (1). Therefore, we have an alternative proof for the result obtained in [15, Theorem 2.1].

Due to Proposition 3.3, *from now on we assume that the sequence $(x^k)_{k\in\mathbb{N}}$ generated by Algorithm 3.1 is infinite.*

3.2 Full asymptotic convergence analysis

The purpose of this section is to prove, under suitable assumptions, the full convergence of the sequence $(x^k)_{k\in\mathbb{N}}$. For this end, we need to be more restrictive with respect to the inexact projection in (3.2) and in the tolerance parameter that controls the non-monotonicity of the line search used in (3.3). More precisely, we assume that in Step 1 of Algorithm 3.1:

A1. For all $k \in \mathbb{N}$, we take $w^k \in \mathcal{R}^{D_k}_{C,\gamma_k}(x^k, z^k)$ with $\gamma_k = (1-\zeta_k)/2$.

It is worth recalling that, taking the parameter $\gamma_k = (1-\zeta_k)/2$, it follows from Lemma 2.14 that $\mathcal{R}^{D_k}_{C,\gamma_k}(x^k, z^k) \subset \mathcal{P}^{D_k}_{C,\zeta_k}(x^k, z^k)$. In addition, we also assume that in Step 2 of Algorithm 3.1:

A2. For all $k \in \mathbb{N}$, we take $0 \leq \nu_k$ such that $\sum_{k=0}^{+\infty} \nu_k < +\infty$.

It follows from Corollary 3.4 that the Armijo and the nonmonotone Average-type line searches discussed in Section 3 satisfy Assumption **A2**.

To prove the full convergence of the sequence $(x^k)_{k\in\mathbb{N}}$ satisfying **A1** and **A2** we consider an additional assumption on the sequence $(D_k)_{k\in\mathbb{N}} \subset \mathcal{D}_\mu$ as follows.

A3. For all $k \in \mathbb{N}$, $(1+\eta_k)D_k - D_{k+1}$ is a positive semidefinite matrix, for some sequence $(\eta_k)_{k\in\mathbb{N}} \subset [0, +\infty)$ such that $\sum_{k\in\mathbb{N}} \eta_k < \infty$.

It is worth mentioning that Assumption **A3** has appeared in the study of the scaled gradient projection method, see, for example, [19]. Note that $D_k = I$ for all $k \in \mathbb{N}$, trivially satisfies **A3**.

We will begin establishing a basic inequality for $(x^k)_{k\in\mathbb{N}}$. To simplify notations, we define the constant

$$\xi := \frac{2\alpha_{\max}}{\sigma} > 0. \tag{3.17}$$?{eq:eta}

Lemma 3.7. *For each $x \in C$ and for all $k \in \mathbb{N}$, we have*

$$\|x^{k+1}-x\|^2_{D_{k+1}} \leq (1+\eta_k)\Big(\|x^k-x\|^2_{D_k} + 2\alpha_k\tau_k\big\langle \nabla f(x^k), x-x^k\big\rangle + \xi\big[f(x^k)-f(x^{k+1})+\nu_k\big]\Big). \tag{3.18}$$?{eq:xkAr

Proof. We know that

$$\|x^{k+1}-x\|^2_{D_k} = \|x^k-x\|^2_{D_k} + \|x^{k+1}-x^k\|^2_{D_k} - 2\langle D_k(x^{k+1}-x^k), x-x^k\rangle,$$

for all $x \in C$ and $k \in \mathbb{N}$. Thus, using (3.4), we have

$$\|x^{k+1}-x\|^2_{D_k} = \|x^k-x\|^2_{D_k} + \tau_k^2\|w^k-x^k\|^2_{D_k} - 2\tau_k\big\langle D_k(w^k-x^k), x-x^k\big\rangle, \qquad \forall\, k \in \mathbb{N}. \tag{3.19}$$?{eq:xkAr

On the other hand, since $w^k \in \mathcal{R}^{D_k}_{C,\gamma_k}(x^k, z^k)$ with $z^k = x^k - \alpha_k D_k^{-1}\nabla f(x^k)$, it follows from Definition 2.10, with $y = x$, $D = D_k$, $u = x^k$, $v = z^k$, $w = w^k$, and $\gamma = \gamma_k$, that

$$\big\langle D_k(x^k - \alpha_k D_k^{-1}\nabla f(x^k) - w^k), x-w^k\big\rangle \leq \gamma_k\|w^k-x^k\|^2_{D_k}, \qquad \forall\, k \in \mathbb{N}.$$

Hence, after some algebraic manipulations in the last inequality, we have

$$-\big\langle D_k(w^k-x^k), x-x^k\big\rangle \leq \alpha_k\big\langle \nabla f(x^k), x-w^k\big\rangle - (1-\gamma_k)\|w^k-x^k\|^2_{D_k}.$$

Combining the last inequality with (3.19), we conclude that

$$\|x^{k+1}-x\|_{D_k}^2 \leq \|x^k-x\|_{D_k}^2 - \tau_k\Big[2(1-\gamma_k)-\tau_k\Big]\|w^k-x^k\|_{D_k}^2 + 2\tau_k\alpha_k\Big\langle \nabla f(x^k), x-w^k\Big\rangle. \quad (3.20)$$?{eq:xkAr

Since $0 \leq \gamma_k < (1-\zeta_{\min})/2 < 1/2$ and $\tau_k \in (0,1]$, we have $2(1-\gamma_k)-\tau_k > \zeta_{\min} > 0$. Thus, it follows from (3.20) that

$$\|x^{k+1}-x\|_{D_k}^2 \leq \|x^k-x\|_{D_k}^2 + 2\tau_k\alpha_k\Big\langle \nabla f(x^k), x-w^k\Big\rangle, \qquad \forall\ k\in\mathbb{N}.$$

Thus, considering that $\Big\langle \nabla f(x^k), x-w^k\Big\rangle = \Big\langle \nabla f(x^k), x-x^k\Big\rangle + \Big\langle \nabla f(x^k), x^k-w^k\Big\rangle$ and taking into account (3.3), we conclude that

$$\|x^{k+1}-x\|_{D_k}^2 \leq \|x^k-x\|_{D_k}^2 + 2\tau_k\alpha_k\Big\langle \nabla f(x^k), x-x^k\Big\rangle + \frac{2\alpha_k}{\sigma}\Big[f(x^k)-f(x^{k+1})+\nu_k\Big], \quad (3.21)$$?{eq;ali}

for all $k\in\mathbb{N}$. On the other hand, applying Lemma 2.18*(iii)* with $x=x^k$, $\alpha=\alpha_k$, $D=D_k$, $w(\alpha)=w^k$, $z=z^k$ and $\zeta=\zeta_k$, we obtain $\langle \nabla f(x^k), w^k-x^k\rangle < 0$, for all $k\in\mathbb{N}$. Therefore, it follows from (3.3) and (3.4) that $0 < -\sigma\tau_k\Big\langle \nabla f(x^k), w^k-x^k\Big\rangle \leq f(x^k)-f(x^{k+1})+\nu_k$, to all $k\in\mathbb{N}$. Hence, due to $0<\alpha_k\leq\alpha_{\max}$, we have

$$\alpha_k[f(x^k)-f(x^{k+1})+\nu_k] < \alpha_{\max}[f(x^k)-f(x^{k+1})+\nu_k], \qquad \forall k\in\mathbb{N}.$$

Therefore, by combining of the last inequality with (3.17) and (3.21) we obtain that

$$\|x^{k+1}-x\|_{D_k}^2 \leq \|x^k-x\|_{D_k}^2 + 2\alpha_k\tau_k\Big\langle \nabla f(x^k), x-x^k\Big\rangle + \xi\Big[f(x^k)-f(x^{k+1})+\nu_k\Big], \quad \forall\ k\in\mathbb{N}.$$

Since **A3** implies that $\|x^{k+1}-x\|_{D_{k+1}}^2 \leq (1+\eta_k)\|x^{k+1}-x\|_{D_k}^2$ the desired inequality (3.18) follows. □

For proceeding with the analysis of the behavior of the sequence $(x^k)_{k\in\mathbb{N}}$, we define the following auxiliary set

$$U := \Big\{x\in C : f(x) \leq \inf_{k\in\mathbb{N}}\Big(f(x^k)+\nu_k\Big)\Big\}.$$

Corollary 3.8. *Assume that f is a convex function. If $U\neq\varnothing$, then $(x^k)_{k\in\mathbb{N}}$ converges to a stationary point of problem (1).*

Proof. Let $x\in U$. Since f is convex, we have

$$0 \geq f(x)-(f(x^k)+\nu_k) \geq \langle \nabla f(x^k), x-x^k\rangle - \nu_k,$$

for all $k\in\mathbb{N}$. Thus, $\langle \nabla f(x^k), x-x^k\rangle \leq \nu_k$, for all $k\in\mathbb{N}$. Using Lemma 3.7, and taking into account that $\tau_k\in(0,1]$ and $0<\alpha_{\min}\leq\alpha_k\leq\alpha_{\max}$, we obtain

$$\|x^{k+1}-x\|_{D_{k+1}}^2 \leq (1+\eta_k)\|x^k-x\|_{D_k}^2 + 2\alpha_{\max}\beta\nu_k + \xi\beta\Big[f(x^k)-f(x^{k+1})+\nu_k\Big], \quad \forall\ k\in\mathbb{N},$$

where $\beta := 1+\sup\{\eta_k :\ k\in\mathbb{N}\}$. Defining

$$\epsilon_k = 2\alpha_{\max}\beta\nu_k + \xi\beta\Big[f(x^k)-f(x^{k+1})+\nu_k\Big],$$

we have

$$\|x^{k+1}-x\|^2_{D_{k+1}} \leq (1+\eta_k)\|x^k-x\|^2_{D_k}+\epsilon_k,$$

for all $k \in \mathbb{N}$. On the other hand, summing ϵ_k with $k = 0, 1, \ldots, N$ and using Corollary 3.4, we have

$$\sum_{k=0}^{N}\epsilon_k \leq 2\alpha_{\max}\beta\sum_{k=0}^{N}\nu_k+\xi\beta\left(f(x^0)-f(x)+\sum_{k=0}^{N+1}\nu_k\right)<+\infty, \qquad \forall N\in\mathbb{N}.$$

Hence, $\sum_{k=0}^{+\infty}\epsilon_k < +\infty$. Thus, it follows from Definition 1.2 that $(x^k)_{k\in\mathbb{N}}$ is quasi-Fejér monotone to U with respect to the sequence $(D_k)_{k\in\mathbb{N}}$. Since U is nonempty, it follows from Theorem 1.4 that $(x^k)_{k\in\mathbb{N}}$ is bounded, and therefore it has cluster points. Let $\bar{x}$ be a cluster point of $(x^k)_{k\in\mathbb{N}}$ and $(x^{k_j})_{j\in\mathbb{N}}$ be a subsequence of $(x^k)_{k\in\mathbb{N}}$ such that $\lim_{j\to\infty} x^{k_j} = \bar{x}$. Considering that f is continuous and $\lim_{k\to+\infty}\nu_k = 0$, we have $\lim_{j\to\infty}(f(x^{k_j})+\nu_{k_j}) = f(\bar{x})$. On the other hand, Lemma 3.2 implies that $\left(f(x^k)+\nu_k\right)_{k\in\mathbb{N}}$ is non-increasing. Thus $\inf_{k\in\mathbb{N}}(f(x^k)+\nu_k) = \lim_{k\to\infty}(f(x^k)+\nu_k) = f(\bar{x})$. Hence, $\bar{x}\in U$, and Theorem 1.4 implies that $(x^k)_{k\in\mathbb{N}}$ converges to $\bar{x}$. The conclusion is obtained by using Proposition 3.3. □

Theorem 3.9. *If f is a convex function and $(x^k)_{k\in\mathbb{N}}$ has no cluster points, then $\Omega^* = \varnothing$, $\lim_{k\to\infty}\|x^k\| = +\infty$, and $\inf_{k\in\mathbb{N}} f(x^k) = \inf\{f(x) : x \in C\}$.*

Proof. Since $(x^k)_{k\in\mathbb{N}}$ has no cluster points, then $\lim_{k\to\infty}\|x^k\| = +\infty$. Assume by contradiction that $\Omega^* \neq \varnothing$. Thus, there exists $\tilde{x}\in C$, such that $f(\tilde{x}) \leq f(x^k)$ for all $k \in \mathbb{N}$. Therefore, $\tilde{x}\in U$. Using Corollary 3.8, we obtain that $(x^k)_{k\in\mathbb{N}}$ is convergent, contradicting that $\lim_{k\to\infty}\|x^k\| = \infty$. Therefore, $\Omega^* = \varnothing$. Now, we claim that $\inf_{k\in\mathbb{N}} f(x^k) = \inf\{f(x) : x \in C\}$. If $\inf_{k\in\mathbb{N}} f(x^k) = -\infty$, the claim holds. Assume by contraction that $\inf_{k\in\mathbb{N}} f(x^k) > \inf_{x\in C} f(x)$. Thus, there exists $\tilde{x}\in C$ such that $f(\tilde{x}) \leq f(x^k) \leq f(x^k)+\nu_k$, for all $k \in \mathbb{N}$. Hence, $U \neq \varnothing$. Using Corollary 3.8, we have that $(x^k)_{k\in\mathbb{N}}$ is convergent, contradicting again $\lim_{k\to\infty}\|x^k\| = +\infty$ and concluding the proof. □

Corollary 3.10. *If f is a convex function and $(x^k)_{k\in\mathbb{N}}$ has at least one cluster point, then $(x^k)_{k\in\mathbb{N}}$ converges to a stationary point of problem (1).*

Proof. Let $\bar{x}$ be a cluster point of the sequence $(x^k)_{k\in\mathbb{N}}$ and $(x^{k_j})_{j\in\mathbb{N}}$ be a subsequence of $(x^k)_{k\in\mathbb{N}}$ such that $\lim_{j\to+\infty} x^{k_j} = \bar{x}$. Considering that f is continuous and $\lim_{k\to+\infty}\nu_k = 0$, we have $\lim_{j\to\infty}(f(x^{k_j})+\nu_{k_j}) = f(\bar{x})$. On the other hand, Corollary 3.4 implies that $(f(x^k)+\nu_k)_{k\in\mathbb{N}}$ is non-increasing. Hence, we have $\inf_{k\in\mathbb{N}}(f(x^k)+\nu_k) = \lim_{k\to\infty}(f(x^k)+\nu_k) = f(\bar{x})$. Therefore $\bar{x}\in U$. Using Corollary 3.8, we obtain that $(x^k)_{k\in\mathbb{N}}$ converges to a stationary point $\tilde{x}\in C$ of problem (1). □

Theorem 3.11. *Assume that f is a convex function and $\Omega^* \neq \varnothing$. Then, $(x^k)_{k\in\mathbb{N}}$ converges to an optimal solution of problem (1).*

Proof. If $\Omega^* \neq \varnothing$, then $U \neq \varnothing$. Therefore, Corollary 3.8 implies that $(x^k)_{k\in\mathbb{N}}$ converges to a stationary point of problem (1) and, due to f be convex, this point is also an optimal solution. □

3.3 Iteration-complexity bound

In the section, we preset some iteration-complexity bounds related to the sequence $(x^k)_{k\in\mathbb{N}}$ generated by Algorithm 3.1. For that, besides assuming **A1** and **A2**, we also need the following assumption.

A4. The gradient ∇f of f is Lipschitz continuous with constant $L > 0$.

For simple notations, we define the following positive constant

$$\tau_{\min} := \min\left\{1, \frac{\underline{\omega}(1-\sigma)}{\alpha_{\max}\mu L}\right\}. \tag{3.22}$$?{eq;taum

Lemma 3.12. *The steepsize τ_k in Algorithm 3.1 satisfies $\tau_k \geq \tau_{\min}$.*

Proof. First, we assume that $\tau_k = 1$. In this case, we have $\tau_k \geq \tau_{\min}$ and the required inequality holds. Now, we assume that $\tau_k < 1$. Thus, it follows from (3.3) that there exists $0 < \hat{\tau}_k \leq \min\{1, \tau_k/\underline{\omega}\}$ such that

$$f\Big(x^k + \hat{\tau}_k(w^k - x^k)\Big) > f(x^k) + \sigma\hat{\tau}_k\Big\langle \nabla f(x^k), w^k - x^k\Big\rangle + \nu_k. \tag{3.23}$$?{eq:ffA1

Considering that we are under assumption **A4**, we apply Lemma 1.1 to obtain

$$f\Big(x^k + \hat{\tau}_k(w^k - x^k)\Big) \leq f(x^k) + \hat{\tau}_k\Big\langle \nabla f(x^k), w^k - x^k\Big\rangle + \frac{L}{2}\hat{\tau}_k^2\|w^k - x^k\|^2. \tag{3.24}$$?{eq:ffA2

Hence, the combination of (3.23) with (3.24) yields

$$(1-\sigma)\Big\langle \nabla f(x^k), w^k - x^k\Big\rangle + \frac{L}{2}\hat{\tau}_k\|w^k - x^k\|^2 > \frac{\nu_k}{\hat{\tau}_k}. \tag{3.25}$$?{eq:ffA3

On the order hand, $w^k \in \mathcal{R}^{D_k}_{C,\gamma_k}(x^k, z^k)$ with $\gamma_k = (1-\zeta_k)/2$, where $z^k = x^k - \alpha_k D_k^{-1}\nabla f(x^k)$. Thus, applying Lemma 2.18*(i)* with $x = x^k$, $w(\alpha) = w^k$, $z = z^k$ and $\zeta = \zeta_k$, we obtain

$$\Big\langle \nabla f(x^k), w^k - x^k\Big\rangle \leq -\frac{1}{2\alpha_k}\|w^k - x^k\|^2_{D_k}.$$

Hence, considering that $\frac{1}{\mu}\|w^k - x^k\|^2 \leq \|w^k - x^k\|^2_{D_k}$ and $0 < \alpha_k \leq \alpha_{\max}$, the last inequality implies

$$\Big\langle \nabla f(x^k), w^k - x^k\Big\rangle \leq -\frac{1}{2\alpha_{\max}\mu}\|w^k - x^k\|^2.$$

The combination of the last inequality with (3.25) yields

$$\left(-\frac{(1-\sigma)}{2\alpha_{\max}\mu} + \frac{L}{2}\hat{\tau}_k\right)\|w^k - x^k\|^2 > \frac{\nu_k}{\hat{\tau}_k} \geq 0.$$

Thus, since $\hat{\tau}_k \leq \tau_k/\underline{\omega}$, we obtain $\tau_k \geq \underline{\omega}\hat{\tau}_k > \underline{\omega}(1-\sigma)/(\alpha_{\max}\mu L) \geq \tau_{\min}$ and the proof is concluded. □

Considering that $\mathcal{R}^{D_k}_{C,\gamma_k}(x^k, z^k) \subset \mathcal{P}^{D_k}_{C,\zeta_k}(x^k, z^k)$, it follows from Lemma 2.18*(ii)* that if $x^k \in \mathcal{R}^{D_k}_{C,\gamma_k}(x^k, z^k)$, then the point x^k is stationary for problem (1). Since $w^k \in \mathcal{R}^{D_k}_{C,\gamma_k}(x^k, z^k)$, the quantity $\|w^k - x^k\|$ can be seen as a measure of stationarity of the point x^k. In next theorem, we present an iteration-complexity bound for this quantity, which is a constrained inexact version of [46, Theorem 1].

Theorem 3.13. *Let $\tau_{\min}$ be defined in (3.22). Then, for every $N \in \mathbb{N}$, the following inequality holds*

$$\min\left\{\|w^k - x^k\| : \ k = 0, 1 \ldots, N-1\right\} \leq \sqrt{\frac{2\alpha_{\max}\mu\left[f(x^0) - f^* + \sum_{k=0}^{\infty}\nu_k\right]}{\sigma\tau_{\min}}}\frac{1}{\sqrt{N}}.$$

Proof. Since $w^k \in \mathcal{R}^{D_k}_{C,\gamma_k}(x^k, z^k)$ with $\gamma_k = (1-\zeta_k)/2$, where $z^k = x^k - \alpha_k D_k^{-1}\nabla f(x^k)$, applying Lemma 2.18*(i)* with $x = x^k$, $w(\alpha) = w^k$, $z = z^k$ and $\zeta = \zeta_k$, and taking into account that $(1/\mu)\|w^k - x^k\|^2 \leq \|w^k - x^k\|^2_{D_k}$ and $0 < \alpha_k \leq \alpha_{\max}$, we obtain

$$\left\langle \nabla f(x^k), w^k - x^k\right\rangle \leq -\frac{1}{2\alpha_k}\|w^k - x^k\|^2_{D_k} \leq -\frac{1}{2\alpha_{\max}\mu}\|w^k - x^k\|^2.$$

The definition of τ_k and (3.3) imply $f(x^{k+1}) - f(x^k) \leq \sigma\tau_k\left\langle \nabla f(x^k), w^k - x^k\right\rangle + \nu_k$. The combination of the last two inequalities together with Lemma 3.12 yields

$$f(x^k) - f(x^{k+1}) + \nu_k \geq \sigma\tau_k\frac{1}{2\alpha_{\max}\mu}\|w^k - x^k\|^2 \geq \sigma\tau_{\min}\frac{1}{2\alpha_{\max}\mu}\|w^k - x^k\|^2.$$

Hence, performing the sum of the above inequality for $k = 0, 1, \ldots, N-1$, we conclude that

$$\sum_{k=0}^{N-1}\|w^k - x^k\|^2 \leq \frac{2\alpha_{\max}\mu\left[f(x^0) - f(x^{N+1}) + \sum_{k=0}^{N}\nu_k\right]}{\sigma\tau_{\min}} \leq \frac{2\alpha_{\max}\mu\left[f(x^0) - f^* + \sum_{k=0}^{\infty}\nu_k\right]}{\sigma\tau_{\min}},$$

which implies the desired result. □ Next we presente some results regarding the number of function evaluations performed by Algorithm 3.1. Note that the computational cost associated to each (outer) iteration involves a gradient evaluation, the computation of a (inexact) projection, and evaluations of f at different trial points. Thus, we must consider the function evaluations at the rejected trial points.

Lemma 3.14. *Let N_k be the number of function evaluations after $k \geq 0$ iterations of Algorithm 3.1. Then, $N_k \leq 1 + (k+1)[\log(\tau_{\min})/\log(\bar{\omega}) + 1]$.*

Proof. Let $j(k) \geq 0$ be the number of inner iterations in Step 2 of Algorithm 3.1 to compute the step size τ_k. Thus, $\tau_k \leq \bar{\omega}^{j(k)}$. Using Lemma 3.12, we have $0 < \tau_{\min} \leq \tau_k$ for all $k \in \mathbb{N}$, which implies that $\log(\tau_{\min}) \leq \log(\tau_k) = j(k)\log(\bar{\omega})$, for all $k \in \mathbb{N}$. Hence, due to $\log(\bar{\omega}) < 0$, we have $j(k) \leq \log(\tau_{\min})/\log(\bar{\omega})$. Therefore,

$$N_k = 1 + \sum_{\ell=0}^{k}(j(\ell) + 1) \leq 1 + \sum_{i=0}^{k}\left(\frac{\log(\tau_{\min})}{\log(\bar{\omega})} + 1\right) = 1 + (k+1)\left(\frac{\log(\tau_{\min})}{\log(\bar{\omega})} + 1\right),$$

where the first equality follows from the definition of N_k. □

Theorem 3.15. *For a given $\epsilon > 0$, Algorithm 3.1 computes x^k and w^k such that $\|w^k - x^k\| \leq \epsilon$ using, at most,*

$$1 + \left(\frac{2\alpha_{\max}\mu\left[f(x^0) - f^* + \sum_{k=0}^{\infty}\nu_k\right]}{\sigma\tau_{\min}}\frac{1}{\epsilon^2} + 1\right)\left(\frac{\log(\tau_{\min})}{\log(\bar{\omega})} + 1\right)$$

function evaluations.

Proof. The proof follows straightforwardly from Theorem 3.13 and Lemma 3.14. □

Theorem 3.16. *Let f be a convex function on C. For a given $\epsilon > 0$, the number of function evaluations performed by Algorithm 3.1 is, at most,*

$$1 + \left(\frac{\|x^0 - x^*\|_{D_0}^2 + \xi \left[f(x^0) - f^* + \sum_{k=0}^{\infty} \nu_k \right]}{2\alpha_{\min}\tau_{\min}} \frac{1}{\epsilon} + 1 \right) \left(\frac{\log(\tau_{\min})}{\log(\bar{\omega})} + 1 \right),$$

to compute x^k such that $f(x^k) - f^* \leq \epsilon$.

Proof. The proof follows straightforwardly from Theorem 3.17 and Lemma 3.14. □ We ended this section with a theorem about iteration-complexity bound for the sequence $\left(f(x^k)\right)_{k\in\mathbb{N}}$ when f is convex.

Theorem 3.17. *Let f be a convex function on C. Assume that the sequence $(D_k)_{k\in\mathbb{N}}$ satisfies* **A3** *with $\eta_k \equiv 0$. Then, for every $N \in \mathbb{N}$,*

$$\min\left\{ f(x^k) - f^* : \ k = 0, 1 \ldots, N-1 \right\} \leq \frac{\|x^0 - x^*\|_{D_0}^2 + \xi \left[f(x^0) - f^* + \sum_{k=0}^{\infty} \nu_k \right]}{2\alpha_{\min}\tau_{\min}} \frac{1}{N}.$$

Proof. Using the first inequality in (3.1) and Lemma 3.12, we have $2\alpha_{\min}\tau_{\min} \leq 2\alpha_k\tau_k$, for all $k \in \mathbb{N}$. We also know form the convexity of f that $\langle \nabla f(x^k), x^* - x^k \rangle \leq f^* - f(x^k)$, for all $k \in \mathbb{N}$. Thus, applying Lemma 3.7 with $x = x^*$ and taking into account that $\eta_k \equiv 0$, after some algebraic manipulations, we conclude

$$2\alpha_{\min}\tau_{\min}\left[f(x^k) - f^* \right] \leq \|x^k - x^*\|_{D_k}^2 - \|x^{k+1} - x^*\|_{D_{k+1}}^2 + \xi \left[f(x^k) - f(x^{k+1}) + \nu_k \right] \quad k = 0, 1, \ldots.$$

Hence, performing the sum of the above inequality for $k = 0, 1, \ldots, N-1$, we obtain

$$2\alpha_{\min}\tau_{\min} \sum_{k=0}^{N-1} \left[f(x^k) - f^* \right] \leq \|x^0 - x^*\|_{D_0}^2 - \|x^{N+1} - x^*\|_{D_N}^2 + \xi \left[f(x^0) - f(x^{N+1}) + \sum_{k=0}^{N-1} \nu_k \right].$$

Thus, $2\alpha_{\min}\tau_{\min} N \min\{ f(x^k) - f^* : \ k = 0, 1 \ldots, N-1 \} \leq \|x^0 - x^*\|_{D_0}^2 + \xi[f(x^0) - f^* + \sum_{k=0}^{N-1} \nu_k]$, which implies the desired inequality. □

Remark 3.18. For suitable choices of the scale matrix and the step size, SGP merges into the well known *spectral gradient method*. Therefore, all obtained results in this section also hold to this method.

Chapter 4

Numerical experiments

This chapter presents some numerical experiments in order to illustrate the potential advantages of considering inexact schemes in the SGP method. We will discuss inexactness associated with both the projection onto the feasible set and the line search procedure.

Given A and B two $m \times n$ matrices, with $m \geq n$, and $c \in \mathbb{R}$, we consider the matrix function $f : \mathbb{R}^{n\times n} \to \mathbb{R}$ given by:

$$f(X) := \frac{1}{2}\|AX - B\|_F^2 + \sum_{i=1}^{n-1} \left[c\left(X_{i+1,i+1} - X_{i,i}^2\right)^2 + (1 - X_{i,i})^2 \right], \tag{4.1}$$?{objfun}

which combines a least squares term with a Rosenbrock-type function. Throughout this section, $X_{i,j}$ stands for the ij-element of the matrix X and $\|\cdot\|_F$ denotes the Frobenius matrix norm, i.e., $\|A\|_F := \sqrt{\langle A, A\rangle}$ where the inner product is given by $\langle A, B\rangle = \mathrm{tr}(A^T B)$. The test problems consist of minimizing f in (4.1) subject to two different feasible sets, as described below. We point out that interesting applications in many areas emerge as constrained least squares matrix problems, see [15] and references therein. In turn, the Rosenbrock term was added in order to make the problems more challenging.

Problem I:

$$\begin{array}{ll} \min & f(X) \\ \text{s.t.} & X \in SDD^+, \\ & L \leq X \leq U, \end{array}$$

where SDD^+ is the cone of symmetric and diagonally dominant real matrices with positive diagonal, i.e.,

$$SDD^+ := \{X \in \mathbb{R}^{n\times n} \mid X = X^T,\ X_{i,i} \geq \sum_{j\neq i} |X_{i,j}|\ \forall i\},$$

L and U are given $n \times n$ matrices, and $L \leq X \leq U$ means that $L_{i,j} \leq X_{i,j} \leq U_{i,j}$ for all i, j. The feasible set of Problem I was considered, for example, in the numerical tests of [15].

Problem II:

$$\begin{array}{ll} \min & f(X) \\ \text{s.t.} & X \in \mathbb{S}_+^n, \\ & \mathrm{tr}(X) = 1, \end{array}$$

where $\mathbb{S}^n_+$ is the cone of symmetric and positive semidefinite real matrices and $\mathrm{tr}(X)$ denotes the trace of X. The feasible set of Problem II was known as *spectrahedron* and appears in several interesting applications see, for example, [4, 43] and references therein.

It is easy to see that the feasible set of Problem I is a closed and convex set and the feasible set of Problem II is a compact and convex set. As discussed in Section 2.4, the Dykstra's alternating algorithm and the Frank-Wolfe algorithm can be used to calculate inexact projections. The choice of the most appropriate method depends on the structure of the feasible set under consideration. For Problem I, we used the Dykstra's algorithm described in [15], see also [63]. In this case, $SDD^+ = \cap_{i=1}^n SDD_i^+$, where

$$SDD_i^+ := \{X \in \mathbb{R}^{n\times n} \mid X = X^T,\ X_{i,i} \geq \sum_{j\neq i} |X_{i,j}|\} \text{ for all } i = 1, \ldots, n,$$

and the projection of a given $Z \in \mathbb{R}^{n\times n}$ onto SDD^+ consists of cycles of projections onto the convex sets SDD_i^+. Here an iteration of the Dykstra's algorithm should be understood as a complete cycle of projections onto all SDD_i^+ sets and onto the box $\{X \in \mathbb{R}^{n\times n} \mid L \leq X \leq U\}$. Recall that this scheme provides an inexact projection as in Definition 2.5. Now consider Problem II. It is well known that calculating an exact projection onto the spectrahedron (i.e., onto the feasible set of Problem II) requires a complete spectral decomposition, which can be prohibitive specially in the large scale case. In contrast, the computational cost of an iteration of the Frank-Wolfe algorithm described in Algorithm 2.1 is associated by an extreme eigenpair computation, see, for example, [51]. Unfortunately, despite its low cost per-iteration, the Frank-Wolfe algorithm suffers from a slow convergence rate. Thus, we considered a variant of the Frank-Wolfe algorithm proposed in [4], which improves the convergence rate and the total time complexity of the classical Frank-Wolfe method. This algorithm specialized for the projection problem over the spectrahedron is carefully described in [1]. Without attempting to go into details, it replaces the top eigenpair computation in Frank-Wolfe with a top-p (with $p \ll n$) eigenpair computation, where p is an algorithmic parameter automatically selected. The total number of computed eigenpairs can be used to measure the computational effort to calculate projections. We recall that a Frank-Wolfe type scheme provides an inexact projection as in Definition 2.10.

We notice that Problems I and II can be seen as particular instances of the problem (1) in which the number of variables is $(n^2 + n)/2$. This mean that they can be solved by using Algorithm 3.1. We are especially interested in the spectral gradient version [15, 14] of the SGP method, which is often associated with large-scale problems [16]. For this, we implemented Algorithm 3.1 considering $D_k := I$ for all k, $\alpha_0 := \min(\alpha_{\max}, \max(\alpha_{\min}, 1/\|\nabla f(x^0)\|))$ and, for $k > 0$,

$$\alpha_k := \begin{cases} \min(\alpha_{\max}, \max(\alpha_{\min}, \langle s^k, s^k\rangle / \langle s^k, y^k\rangle)), & \text{if } \langle s^k, y^k\rangle > 0 \\ \alpha_{\max}, & \text{otherwise,} \end{cases}$$

where $s^k := X^k - X^{k-1}$, $y^k := \nabla f(X^k) - \nabla f(X^{k-1})$, $\alpha_{\min} = 10^{-10}$, and $\alpha_{\max} = 10^{10}$. We set $\sigma = 10^{-4}$, $\underline{\tau} = 0.1$, $\bar{\tau} = 0.9$, $\mu = 1$ and $\nu_0 = 0$. Parameter $\delta_{\min}$ was chosen according to the

line search used (see Section 3), while parameter $\zeta_{\min}$ depends on the inexact projection scheme considered.

In the line search scheme (Step 2 of Algorithm 3.1), if a step size τ_{trial} is not accepted, then τ_{new} is calculated using one-dimensional quadratic interpolation employing the safeguard $\tau_{\text{new}} \leftarrow \tau_{\text{trial}}/2$ when the minimum of the one-dimensional quadratic lies outside $[\underline{\omega}\tau_{\text{trial}}, \bar{\omega}\tau_{\text{trial}}]$, see, for example, [58, Section 3.5]. Concerning the stopping criterion, all runs were stopped at an iterate X^k declaring convergence if

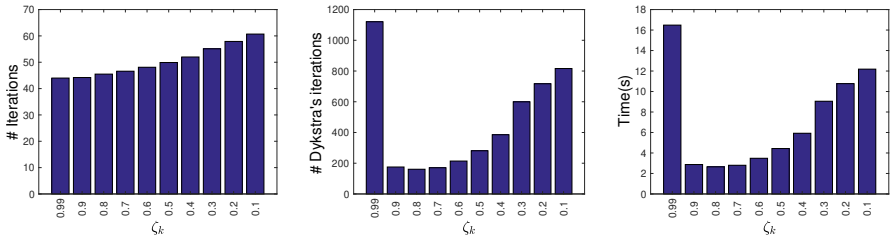
$$\max_{i,j}(|X^k_{i,j} - W^k_{i,j}|) \leq 10^{-6},$$

where W^k is as in (3.2). Our codes are written in Matlab and are freely available at `https://github.com/maxlemes/SGP`. All experiments were run on a macOS 10.15.7 with 3.7GHz Intel Core i5 processor and 8GB of RAM.

4.1 Influence of the inexact projection

We begin the numerical experiments by checking the influence of the forcing parameters that control the degree of inexactness of the projections in the performance of the method. In this first battery of tests, we used Armijo line searches, see Section 3.

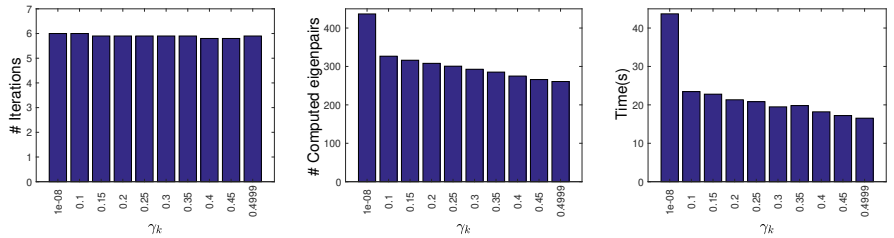
We generated 10 instances of Problem I using $n = 100$, $m = 200$, and $c = 10$. The matrices A and B were randomly generated with elements belonging to $[-1, 1]$. We set $L \equiv 0$ and $U \equiv \infty$ as in [15]. For each instance, the starting point X^0 was randomly generated with elements belonging to $[0, 1]$, then it was redefined as $(X^0 + (X^0)^T)/2$ and its diagonal elements were again redefined as $2\sum_{j\neq i}^{n} X_{i,j}$, ensuring a feasible starting point. Figure 4.1 shows the average number of iterations, the average number of Dykstra's iterations, and the average CPU time in seconds needed to reach the solution for different choices of ζ_k, namely, $\zeta_k = 0.99$, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, and 0.1 for all k. Remember that *smaller values* of ζ_k imply *more inexact* projections. As expected, the number of iterations tended to increase as ζ_k decreased, see Figure 4.1(a). On the other hand, the computational cost of an outer iteration (which can be measured by the number of Dykstra's iterations) tends to decrease when considering smaller values of ζ_k. This suggests a trade-off, controlled by parameter ζ_k, between the number and the cost per iteration. Figure 4.1(b) shows that values for ζ_k close to 0.8 showed better results, which is in line with the experiments reported in [15]. Finally, as can be seen in Figure 4.1(c), the CPU time was shown to be directly proportional to the number Dykstra's iterations.



(a) (b) (c)

Figure 4.1: Results for 10 instances of Problem I using $n = 100$, $m = 200$, and $c = 10$. Average number of: (a) iterations; (b) Dykstra's iterations; (c) CPU time in seconds needed to reach the solution for different choices of ζ_k.

Although Algorithm 3.1 is given only in terms of parameter ζ_k, we will directly consider parameter γ_k for Problem II in which inexact projections are computed according to Definition 2.10. We randomly generated 10 instances of Problem II with $n = 800$, $m = 1000$, and $c = 100$. Matrices A and B were obtained similarly to Problem I. In turn, a starting point X^0 was randomly generated with elements in the interval $[-1, 1]$, then it was redefined to be $X^0(X^0)^T/\text{tr}(X^0(X^0)^T)$, resulting in a feasible initial guess. Figure 4.2 shows the average number of iterations, the average number of computed eigenpairs, and the average CPU time in seconds needed to reach the solution for different constant choices of γ_k ranging from 10^{-8} to 0.4999. Now, *higher values* of γ_k imply *more inexact* projections. Note that for appropriate choices of ζ_k, the adopted values of γ_k fulfill Assumption A1 of Section **??**. Concerning the number of iterations, as can be seen in Figure 4.2(a), the algorithm was not very sensitive to the choice of parameter γ_k. Hence, since higher values of γ_k imply cheaper iterations, the number of computed eigenpairs and the CPU time showed to be inversely proportional to γ_k, see Figures 4.2(b)–(c). Thus, our experiments suggest that the best value for γ_k seems to be 0.4999.



(a) (b) (c)

Figure 4.2: Results for 10 instances of Problem II using $n = 800$, $m = 1000$, and $c = 100$. Average number of: (a) iterations; (b) computed eigenpairs; (c) CPU time in seconds needed to reach the solution for different choices of γ_k.

4.2 Influence of the line search scheme

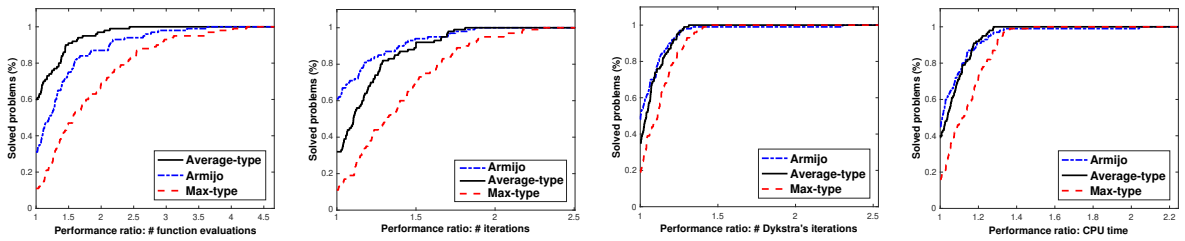
The following experiments compare the performance of the algorithm with different strategies for computing the step sizes. We considered the Armijo, the Average-type, and the Max-type line searches discussed in Section 3. Based on our numerical experience, we employed the fixed value

$\eta_k = 0.85$ for the Average-type line search and $M = 5$ for the Max-type line search. According to the results of the previous section, we used the fixed forcing parameters $\zeta_k = 0.8$ and $\gamma_k = 0.4999$ to compute inexact projections for Problems I and II, respectively.

We randomly generated 100 instances of each problem as described in Section 4.1. The dimension of the problems and the parameter c in (4.1) were also taken arbitrarily. For Problem I, we choose $100 \leq n \leq 800$ and $10 \leq c \leq 50$, whereas for Problem II, we choose $10 \leq n \leq 200$ and $100 \leq c \leq 1000$. In both cases, we set $m = 2n$. We compare the strategies with respect to the number of function evaluations, the number of (outer) iterations, the total computational effort to calculate projections (measured by the number of Dykstra's iterations and computed eigenpairs for Problems I and II, respectively), and the CPU time. The results are shown in Figures 4.3 and 4.4 for Problems I and II, respectively, using performance profiles [33].

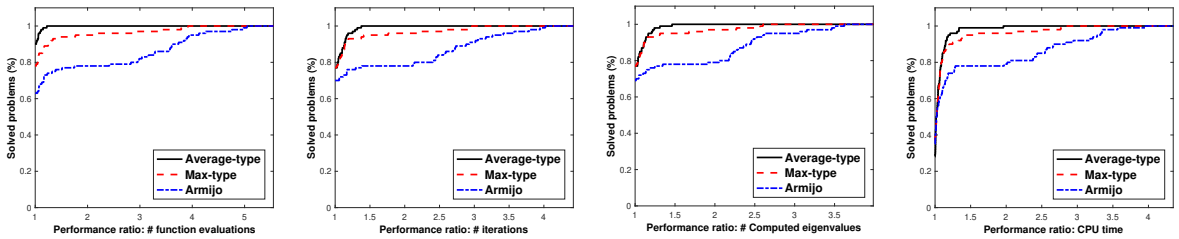
For Problem I, with regard to the number of function evaluations, the algorithm with the Average-type line search was the most efficient among the tested strategies. In a somewhat surprising way, in this set of test problems, the Armijo strategy was better than the Max-type line search, see Figure 4.3(a). On the other hand, as can be seen in Figure 4.3(b), the Armijo strategy required fewer iterations than the nonmonotone strategies. As expected, this was reflected in the number of Dykstra's iterations and the CPU time, see Figures 4.3(c)–(d). We can conclude that, with respect to the last two criteria, the Armijo and Average-type strategies had similar and superior performances to the Max-type strategy.

Now, concerning Problem II, Figure 4.4 shows that the nonmonotone strategies outperformed the Armijo strategy in all the comparative criteria considered. Again, the Average-type strategy seems to be superior to the Max-type strategy.



(a) (b) (c) (d)

Figure 4.3: Performance profiles for Problem I considering Algorithm 3.1 with the Armijo, the Average-type, and the Max-type line searches strategies using as performance measurement: (a) number of function evaluations; (b) number of (outer) iterations; (c) number of Dykstra's iterations; (d) CPU time.



(a) (b) (c) (d)

Figure 4.4: Performance profiles for Problem II considering Algorithm 3.1 with the Armijo, the Average-type, and the Max-type line searches strategies using as performance measurement: (a) number of function evaluations; (b) number of (outer) iterations; (c) number of computed eigenpairs; (d) CPU time.

From all the above experiments, we conclude that the nonmonotone line searches tend to require fewer objective function evaluations. However, this does not necessarily mean computational savings, since there may be an increase in the number of iterations. In this case, optimal efficiency of the algorithm comes from a compromise between those two conflicting tendencies. Overall, the use of nonmonotone line search techniques is mainly justified when the computational effort of an iteration is associated with the cost of evaluating the objective function.

Chapter 5

Applications

5.1 Risks measures

In this section we discuss different ways to measure the risk of portfolio.

Definition 5.1. *A portfolio with n assets, $A_1, A_2, \ldots, A_n$, is a vector $x^\top = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ where each coordinate x_i is the weight of capital invested in asset A_i.*

Let $L(x, y)$ be the loss associated with the decision vector x, to be chosen from a certain subset X of $\mathbb{R}^n$ and the random vector $y \in \mathbb{R}^m$. The vector x can be interpreted as representing a portfolio, with X as the set of available portfolios (subject to various constraints). The vector y stands for the uncertainties, e.g. in market parameters, that can affect the loss. Of course the loss might be negative and thus, in effect, constitute a gain.

For each x, the loss $Y = L(x, y)$ is a random variable having a distribution in $\mathbb{R}$ induced by that of y. The underlying probability distribution of $y \in \mathbb{R}^m$ will be assumed to have density, which we denote by $p(y)$.

The probability of $L(x, y)$ not exceeding a threshold α is given by

$$\Psi(x, \alpha) = \int_{L(x,y)\leq\alpha} p(y)\, dy.$$

As a function of α for fixed x, $\Psi(x, \alpha)$ is the *cumulative distribution function* for the loss associated with x. Since $p(y)$ is continuous, $\Psi(x, \alpha)$ is nondecreasing with respect to α and continuous.

The probability of loss is a risk measure of a portfolio. However, in 1999 Artzner, Delbaen, Eber and Heath (see in [6]) defined, what they called *coherent risk measure*, from the following axioms to a risk measure $\mathcal{R}$:

Axiom 1 (Translation invariance)**.** *For all portfolio x and all $m \in \mathbb{R}$, we have*

$$\mathcal{R}(x + m) = \mathcal{R}(x) - m.$$

This means that adding an amount of cash, m, to a portfolio decreases your risk by the same amount.

Axiom 2 (Subadditivity). *If x_A and x_B are two portfolios, then*

$$\mathcal{R}(x_A + x_B) \leq \mathcal{R}(x_A) + \mathcal{R}(x_B),$$

In particular, this property means that "a merger does not create extra risk."

Axiom 3 (Positive homogeneity). *For all portfolio x and all $\lambda \leq 0$ we have*

$$\mathcal{R}(\lambda x) = \lambda \mathcal{R}(x), \qquad se\ \lambda \geq 0.$$

This property implies that the risk measure is a linear function of the size of the position.

Axiom 4 (Monotonicity). *Let x_A, x_B be two portfolios,*

$$if\ x_A \prec x_B, \quad then \quad \mathcal{R}(x_A) \geq \mathcal{R}(x_B).$$

If a portfolio x_A performs worse than x_B in any scenario ($x_A \prec x_B$), then it means that the portfolio x_A is riskier than x_B.

Definition 5.2. *A risk measure satisfying the axioms of translation invariance, subadditivity, positive homogeneity, and monotonicity is called* ***coherent****.*

When the risk measure is not assumed to have variations proportional to the risk variations themselves, the positive homogeneity is no longer satisfied. Alternative axioms can be proposed.

Axiom 5 (Convexity). *For all portfolios x_A, x_B and for all $0 \leq \lambda \leq 1$,*

$$\mathcal{R}(\lambda x_A + (1-\lambda)x_A) \leq \lambda \mathcal{R}(x_A) + (1-\lambda)\mathcal{R}(x_B).$$

This axiom was proposed in 2002 by Föllmer and Shield (see in [39]). It means that diversification does not increase risk.

Definition 5.3. *A risk measure satisfying the axioms of translation invariance, monotonicity, and convexity is called* ***convex****.*

Proposition 5.4. *A convex risk measure is coherent if it satisfies the positive homogeneity. Note also that the positive homogeneity and the subadditivity implies the convexity.*

In financial business, some of the risk management requirements are done in terms of loss distribution percentiles. An upper percentile of the loss distribution is called *Value-at-Risk* (VaR_β)[1]. For instance, $\text{VaR}_{0.95}$ is an upper estimate of losses which is exceeded with 5% probability. The popularity of VaR_β is mostly related to a simple and easy to understand representation of high losses. VaR_β can be quite efficiently estimated and managed when underlying risk factors are normally (log-normally) distributed.

[1]By definition, VaR_β is the percentile of the loss distribution, i.e., with a specified confidence level β, the VaR_β of a portfolio is the lowest amount ζ such that, with probability β, the loss is less or equal to ζ.

An alternatuve measure of losses is another percentile risk measure which is called *Conditional Value-at-Risk* (CVaR_β). The CVaR_β risk measure is closely related to VaR_β. For continuous distributions, CVaR_β is defined as the conditional expected loss under the condition that it exceeds VaR_β, see Rockafellar and Uryasev (see [64]). For continuous distributions, this risk measure also is known as *Expected Shortfall*, *Mean Excess Loss*, *Mean Shortfall*, or *Tail Value-at-Risk*.

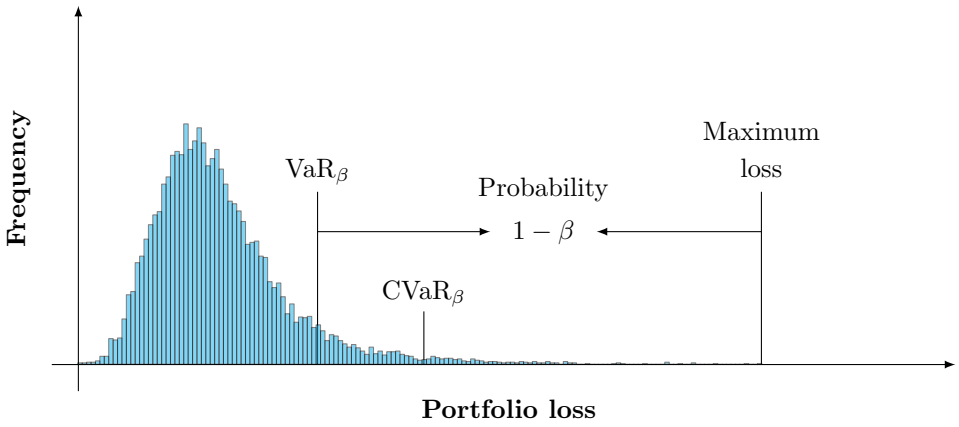


Figure 5.1: Portfolio Loss Distribution, VaR and CVaR

Remark 5.5.

- The VaR_β answers the question: what is the maximum loss with a specified confidence level β?
- The CVaR_β answers the following question: what is the average loss in the $100(1-\beta)\%$ worst case scenarios?

The VaR_β and CVaR_β values for the loss random variable $L(x,y)$ associated with x and any specified probability level $\beta \in (0,1)$ will be denoted by $\alpha_\beta(x)$ and $\text{ES}_\beta(x)$. In our setting they are given by

$$\alpha_\beta(x) = \min\{\alpha \in \mathbb{R} : \Psi(x,\alpha) \geq \beta\}, \tag{5.1}$$?{eq:var}

and

$$\text{ES}_\beta(x) = \frac{1}{1-\beta} \int_{L(x,y)\geq \alpha_\beta(x)} L(x,y)\, p(y)\, dy \tag{5.2}$$?{eq:cvar

Although VaR_β is a very popular measure of risk, VaR_β is not a coherent risk measure as it does not satisfy subadditivity, but CVaR is a coherent risk measure.

Let $F_\beta : X \times \mathbb{R} \to \mathbb{R}$ be the function defined by

$$F_\beta(x,\alpha) = \alpha + \frac{1}{1-\beta} \int_{y\in\mathbb{R}^m} [L(x,y)-\alpha]^+ \, p(y)\, dy,$$

where $[t]^+ = \max t, 0$. The crucial features of F_α, under the assumptions made above, are as follows [64, Theorem 2].

Theorem 5.6. *Minimizing the $CVar_\beta$ of the loss associated with x over all $x \in X$ is equivalent to minimizing $F_\beta(x, \alpha)$ over all $(x, \alpha) \in X \times \mathbb{R}$, in the sense that*

$$\min_{x \in X} ES_\beta(x) = \min_{(x,\alpha) \in X \times \mathbb{R}} F_\beta(x, \alpha),$$

where moreover a pair (x^*, α^*) achieves the right hand side minimum if and only if x^* achieves the left hand side minimum and $\alpha^* \in A_\beta(x^*)$, where

$$A_\beta(x^*) = \operatorname*{argmin}_{\alpha \in \mathbb{R}} F_\beta(x, \alpha).$$

In particular, therefore, in circumstances where the interval $A_\beta(x^*)$ reduces to a single point (as is typical), the minimization of $F_\beta(x, \alpha)$ over $(x, \alpha) \in X \times \mathbb{R}$ produces a pair (x^*, α^*), not necessarily unique, such that x^* minimizes the $CVaR_\beta$ and α^* gives the corresponding VaR_β.

Furthermore, $F_\beta(x, \alpha)$ is convex with respect to (x, α) and $ES_\beta(x)$ is convex with respect to x, when $L(x, y)$ is convex with respect to x, in which case, if the constraints are such that X is a convex set, the join minimization is an instance of convex programming.

5.2 The Gaussian Case

We now consider the case where the decision vector x represents a portfolio of financial instruments in the sense that $x^\top = (x_1, x_2, \ldots, x_n)$, with x_i being the position in instrument i and denoting by y_i the return on instrument i, we take the random vector to be $y^\top = (y_1, y_2, \ldots, y_n)$. The distribution of y constitutes a joint distribution of the various returns and is independent of x; it has density $p(y)$.

The return on a portfolio x is the sum of the returns on the individual instruments in the portfolio, scaled by the proportions x_i. The loss, being the negative of this, is therefore given by

$$L(x, y) = -[x_1 y_1 + x_2 y_2 + \cdots x_n y_n] = -x^\top y.$$

As long as $p(y)$ is continuous with respect to y, the cumulative distribution function for the loss associated with x will itself be continuous and although VaR and CVaR are usually defined in terms of monetary value, they could be defined as percentage returns the same happens to loss function.

For a closer look, let $\mu(x)$ and $\sigma(x)$ denote the mean and variance of the loss associated with portfolio x; in terms of the mean μ and variance Σ of y, we have

$$\mu(x) = -x^\top \mu \quad \text{and} \quad \sigma^2(x) = x^\top \Sigma x.$$

We assume now that the random variable of loss $Y = L(x, y)$ is normally distributed, that is, $Y \sim N(\mu, \sigma^2)$. Let $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ be the probability density function of Standard Normal Distribution and $\Phi(x) = \int_{-\infty}^{x} \phi(z) dz$ the standard normal accumulated density function.

Since $P\Big[L(x,y) \leq \alpha\Big] = \Psi(x,\alpha)$, where $P\Big[L(x,y) \leq \alpha\Big]$ is the probability that loss is less than α. By definition we have $P\Big[L(x,y) \leq \alpha_\beta(x)\Big] = \beta$, which standardizing we obtain:

$$P\left[\frac{L(x,y)-\mu(x)}{\sigma(x)} \leq \frac{\alpha_\beta(x)-\mu(x)}{\sigma(x)}\right] = \beta.$$

Hence,

$$\frac{\alpha_\beta(x)-\mu(x)}{\sigma(x)} = \Phi^{-1}(\beta)$$

and then

$$\alpha_\beta(x) = \mu(x) + \Phi^{-1}(\beta)\sigma(x). \tag{5.3}$$?{eq:var2

This is a special case of the standard deviation-based risk measure with $c = \Phi^{-1}(\beta)$. It implies that the value-at-risk is a coherent and convex risk measure if the asset returns are normally distributed. The expression of the expected shortfall is:

$$\mathrm{ES}_\beta(x) = \frac{1}{1-\beta}\int_{\alpha_\beta(x)}^{\infty} y\, p(y)\, dy,$$

considering $p(y)$ the normal density function, we have that:

$$\mathrm{ES}_\beta(x) = \frac{1}{1-\beta}\int_{-\mu(x)+\sigma(x)\Phi^{-1}(\beta)}^{\infty} \frac{y}{\sigma(x)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y+\mu(x)}{\sigma(x)}\right)^2} dy,$$

With the variable change $t = \frac{y+\sigma(x)}{\sigma(x)}$, we obtain

$$\begin{aligned}
\mathrm{ES}_\beta(x) &= \frac{1}{1-\beta}\int_{\Phi^{-1}(\beta)}^{\infty} (-\mu(x)+\sigma(x)t)\frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \\
&= \frac{\mu(x)}{1-\beta}[\Phi(t)]_{\Phi^{-1}(\beta)}^{\infty} + \frac{\sigma(x)}{(1-\beta)\sqrt{2\pi}}\int_{\Phi^{-1}(\beta)}^{\infty} t e^{-t^2/2} dt \\
&= \mu(x) + \frac{\sigma(x)}{(1-\beta)\sqrt{2\pi}}\Big[-e^{-t^2/2}\Big]_{\Phi^{-1}(\beta)}^{\infty} \\
&= \mu(x) + \frac{\sigma(x)}{(1-\beta)\sqrt{2\pi}} e^{-\frac{[\Phi^{-1}(\beta)]^2}{2}}
\end{aligned}.$$

So CVaR can be calculated by

$$\mathrm{ES}_\beta(x) = \mu(x) + \frac{\phi(\Phi^{-1}(\beta))}{1-\beta}\sigma(x). \tag{5.4}$$?{eq:cvar

Like the value-at-risk, it is a standard deviation-based risk measure with $c = \phi(\Phi^{-1}(\beta))/(1-\beta)$ In the Gaussian world, different risk measures can be calculated using the expected return and volatility.

The volatility of the loss, $\sigma^2(x)$, was defined as a risk measure to a portfolio by Nobel laureate in economics, Harry Markowitz in 1952 (see in [54]) under the hypothesis of loss has a normal distribution. Note that, in this case, by (5.3) and (5.4), that both Var and CVaR are the form $\mu(x) + c\sigma(x)$. In general, we want a portfolio with positive returns, that is, $\mu(x) \leq 0$. If the portfolio manager has very optimistic forecasts, component $\mu(x)$ may substantially reduce the

risk measure. This explains why omitting the mean component is standard practice in the asset management industry. Under the hypothesis of loss has a normal distribution volatility, VaR and CVaR are coherent risk measures.

Example 5.7. We consider three stocks A, B and C whose current prices are respectively \$15.00, \$25.00 and \$30.00. We assume that their expected returns are equal to 30 bps[2], 50 bps and 20 bps on a daily basis and their daily volatilities are 3%, 2%, and 1% respectively. The asset correlation matrix is given by

$$\rho = \begin{pmatrix} 1.00 & 0.40 & 0.15 \\ 0.40 & 1.00 & 0.60 \\ 0.15 & 0.60 & 1.00 \end{pmatrix}$$

We consider a portfolio composed by 100 stocks A, 200 stocks B and 100 stocks C. The value of this portfolio is $\$9,500.00$, being $\$1,500.00$ invested in stock A, $\$5,000.00$ in stock B and $\$3,000.00$ in stock C, therefore the weights of the stocks in this portfolio are 15.79%, 52.63%, and 31.58% respectively, so $x = (0.1579;\ 0.5263;\ 0.3158)$. The expected loss of the portfolio is $\mu(x) = -30 \times 0.1579 + 50 \times 0.5263 + 20 \times 0.3158$, that is, $\mu(x) = -37$ bps. Using the relationship $\Sigma_{i,j} = \rho_{i,j}\sigma_1\sigma_j$ we find the covariance matrix

$$\Sigma = \begin{pmatrix} 9.0 & 2.4 & 4.5 \\ 2.4 & 4.0 & 1.2 \\ 4.5 & 1.2 & 1.0 \end{pmatrix} \times 10^{-4},$$

since the volatility of portfolio is given by $\sigma^2(x) = x^\top \Sigma x$, we have that $\sigma(x) = 1.51\%$.

Considering $\alpha = 0.99$, we have $\Phi^{-1}(0.99) = 2.325$ and $\phi(\Phi^{-1}(0.99)) = \phi(2.325) = 2.68\%$. Using the equations (5.3) and (5.4) we have

$$\begin{aligned} \mathrm{VaR}_{99\%}(x) &= -0.37\% + 2.325 \times 1.51\% = 3.14\% \\ \mathrm{ES}_{99\%}(x) &= -0.37\% + \frac{2.68}{0.01} \times 1.51\% = 3.68\%. \end{aligned}$$

Risk can also be expressed in monetary terms, in this case,

$$\begin{aligned} \mathrm{VaR}_{99\%}(x) &= 3.14\% \times \$9,500.00 = \$298.30 \\ \mathrm{ES}_{99\%}(x) &= 3.68\% \times \$9,500.00 = \$349.60. \end{aligned}$$

According to the result of VaR, in 99% of the days, the loss will be less than \$298.30, however in the 1% of the remaining days the VaR does not measure how much great can be the loss, this is the role of CVaR which indicates that the average loss will be \$349.60 on the worst 1% days.

5.3 Portfolio optimization

We consider a universe of n assets. Let $x^\top = (x_1, x_2, \ldots, x_n)$ be a portfolio and denoting by R_i the return of asset i. If $R^\top = (R_1, R_2, \ldots, R_n)$ is the random vector of asset returns then the return

[2] 1 bps or one basis point is equivalent to 0.01% (the hundredth part of 1%) or 0.0001 in decimal form.

of portfolio x is given by:

$$R(x) = [x_1 R_1 + x_2 R_2 + \cdots x_n R_n] = -x^\top R.$$

If we note μ and Σ the vector of expected returns and the covariance matrix of asset returns, we deduce that the expected return of portfolio x is equal to:

$$\mu(x) = -x^\top \mu$$

whereas its variance is given by:

$$\sigma^2(x) = x^\top \Sigma x.$$

The mean-variance portfolio (MVP) of Markowitz ([54]) consists in maximizing the expected return $\mu(x)$ for a fixed value σ_0 of the volatility $\sigma(x)$. This can be achieved by solving a standard quadratic programming (QP) problem:

$$\begin{aligned} &\min_{x \in X} \; -\mu(x), \\ \text{s.t.} \; &\begin{cases} \sigma(x) = \sigma_0, \\ \mathbf{1}^\top x = 1 \end{cases} \end{aligned} \tag{5.5}$$ {?}

where $\mathbf{1}^\top = (1, 1, \ldots, 1)$. The set of constraints X could be defined by one or more properties, for example:

- Capital constraint: $\mathbf{1}^\top x = 1$;
- Long-only constraint: $x \geq 0$;
- Self-financial constraint: $\mathbf{1}^\top x = 0$;
- Holding constraint: $I \leq x \leq W$ where $I, W \in \mathbb{R}^n$ are lower and upper bounds of the asset positions, respectively.
- Leverage constraint: $\|x\|_1 \leq K$.

The portfolio optimization is defined when we want to minimize (or maximize) a performance measure subject to a set of constraints. To performance measures there are, for example:

- Expected return: $\mu(x)$;
- Volatility: $\sigma(x)$;
- Sharpe Ratio (SR): expected return per unit of risk

  $$SR = \frac{x^\top \mu - r_f}{\sigma(x)}$$

  where r_f is the risk-free rate (e.g. interest rate on a Treasury bill);

- Information Ratio (IR): Sharpe Ratio with $r_f = 0$;
- VaR (Value at Risk): quantile of the loss;
- CVaR (Conditional Value at Risk): expected value of the loss above some quantile.

We can choose a lot of optimization problems, like global minimum CVaR portfolio (GMCP)

$$\min_{x \in X} \; ES_\beta(x), \qquad (5.6)\ \{?\}$$
$$\text{s.t.} \begin{cases} \mathbf{1}^\top x = 1 \\ \mathbf{0} \leq x \leq \mathbf{1} \end{cases}$$

or

$$\min_{x \in X} \; \lambda ES_\beta(x) - \mu(x), \qquad (5.7)\ \{?\}$$
$$\text{s.t.} \begin{cases} \mathbf{1}^\top x = 1 \\ \mathbf{0} \leq x \leq \mathbf{1} \end{cases}$$

where λ is a parameter that controls how risk-averse the investor is.

5.4 Risk Parity

Markowitz's portfolio has never been fully embraced by practitioners, among other reasons because it only considers the risk of the portfolio as a whole and ignores the risk diversification (i.e., concentrates risk too much in few assets, this was observed in the 2008 financial crisis): one solution is the risk parity portfolio.

Risk parity is an approach to portfolio management that focuses on allocation of risk rather than allocation of capital. The risk parity approach asserts that when asset allocations are adjusted to the same risk level, the portfolio can achieve a higher Sharpe ratio and can be more resistant to market downturns.

While the minimum variance portfolio tries to minimize the variance (with the disadvantage that a few assets may be the ones contributing most to the risk), the risk parity portfolio tries to constrain each asset (or asset class, such as bonds, stocks, real estate, etc.) to contribute equally to the portfolio overall volatility.

To define a risk-based allocation strategies it is necessary define how the risk of an asset affects the risk of the portfolio. Let $x^\top = (x_1, x_2, \ldots, x_n)$ a portfolio with n assets and $\mathcal{R}(x)$ be a coherent risk measure of x. Since $\mathcal{R}(x)$ is homogeneous, we have

$$\mathcal{R}(x) = \frac{d}{d\lambda}\mathcal{R}(\lambda x) = \sum_{i=1}^{n} x_i \frac{\partial \mathcal{R}(x)}{\partial x_i}$$

we can define the risk contribution of the asset i by $\mathcal{RC}_i$, where

$$\mathcal{RC}_i(x) = x_i \frac{\partial \mathcal{R}(x)}{\partial x_i}.$$

Therefore, the risk can be written as follows

$$\mathcal{R}(x) = \sum_{i=1}^{n} \mathcal{RC}_i(x)$$

which is known as **Euler's Allocation Principle**.

In the Gaussian case, that is, assuming that asset returns have a Normal Distribution, we can consider $\mathcal{R}(x) = \sigma(x) = \sqrt{x^\top \Sigma x}$, it follows that

$$\mathcal{RC}_i(x) = x_i \frac{\partial \sigma(x)}{\partial x_i} = \frac{x_i (\Sigma x)_i}{\sqrt{x^\top \Sigma x}}.$$

It may be interesting for the investor to choose different risks for different assets, for example, the investor may want to take a greater (or lesser) risk for assets in a certain sector. The investor then chooses the percentage of risk that each asset should have in the portfolio, that is,

$$\mathcal{RC}_i(x) = b_i \mathcal{R}(x),$$

where b_i represents the risk contribution that asset i will have in the portfolio. Of course $b_i \geq 0$ for all i and $\mathbf{1}^\top b = 1$, with $b = (b_1, b_2, \ldots, b_n)$. We can see that a PDR reduces to a PPR when $b_i = 1/n$, for all i.

In general, the asset allocation problem consists of determining the weight x_i that each asset will have in the portfolio in order to follow the chosen strategy. There is no challenge in PPI allocation, but to allocate the assets in a PDR we need to solve the following non-linear system

$$\begin{cases} \mathcal{RC}_i(x) = b_i \mathcal{R}(x), \\ b_i \geq 0, \\ x_i \geq 0, \\ \mathbf{1}^\top b = 1, \\ \mathbf{1}^\top x = 1. \end{cases} \tag{5.8}$$?{eq:SisN

Only in trivial cases can we find analytical solutions for this system, however, we can always find numerical solutions.

Consider $b^\top = (b_1, b_2, \ldots, b_n)$, with $\mathbf{1}^\top b = 1$ and $f(x; b)$ defined by:

$$f(x; b) = \sum_{i=1}^{n} \Big(\mathcal{RC}_i(x) - b_i \mathcal{R}(x) \Big)^2$$

and the following optimization problem

$$\begin{aligned} &\min_{x \geq \mathbf{0}} f(x; b); \\ &\text{Subject to conditions: } \mathbf{1}^\top x = 1 \text{ and } \mathbf{0} < x < \mathbf{1}. \end{aligned} \tag{5.9}$$?{eq:MinP

If $x^\star$ is the solution to the problem (5.9) and $f(x^\star; b) = 0$, so $x^\star$ is also system solution (5.8). So we can use some optimization algorithm quadratic to find a solution to the PDR. as it was said before, to find a PPR just choose $b_i = 1/n$, in the process of choosing a PDR.

In the Gaussian case, when asset returns have a Normal distribution, we can consider $\mathcal{R}(x) = \sigma(x)$ and the system (5.8) reduces to

$$\begin{cases} x_i(\Sigma x)_i = b_i x^\top \Sigma x, \qquad i = 1, 2, \ldots, n., \\ b_i \geq 0, \\ x_i \geq 0, \\ \mathbf{1}^\top b = 1, \\ \mathbf{1}^\top x = 1. \end{cases} \tag{5.10}$$?{eq:SisG

Setting $w = x/\sqrt{x^\top \Sigma x}$, the equation $x_i(\Sigma x)_i = b_i x^\top \Sigma x$ is equivalent to $w_i(\Sigma w)_i = b_i$, or, in vector form

$$\Sigma w = b/w.$$

In 2013, Spinu (see in [67]) showed that the convex function

$$f(w) = \frac{1}{2} w^\top \Sigma w - b^\top log(w)$$

has a gradient equal to

$$\nabla f(w) = \Sigma w - b/w$$

and the system (5.10) can be reformulated by the following optimization problem

$$\begin{array}{ll} \min\limits_{x \geq \mathbf{0}} & \frac{1}{2} w^\top \Sigma w - b^\top \log(w); \\ \text{Subject to conditions:} & \mathbf{1}^\top x = 1 \text{ and } \mathbf{0} < x < \mathbf{1}, \end{array} \tag{5.11}$$?{eq:Gaus

whose optimality condition is $\nabla f(w) = 0$ or $\Sigma w = b/w$ which is precisely the solution of the system (5.10).

5.5 Case Study

Chapter 6

Conclusions

In this paper, we study the SGP method to solve constrained convex optimization problems employing inexact projections onto the feasible set and a general nonmonotone line search. We expect that this paper will contribute to the development of research in this field, mainly to solve large-scale problems when the computational effort of an iteration is associated with the projections onto the feasible set and the cost of evaluating the objective function. Indeed, the idea of using the inexactness in the projection as well as in the line search, instead of the exact ones, is particularly interesting from a computational point of view. In particular, it is noteworthy that the Frank-Wolfe method has a low computational cost per iteration resulting in high computational performance in different classes of compact sets, see [40, 51]. An issue that deserves attention is the search for new efficient methods such as the Frank-Wolfe's and Dykstra's methods that generate inexact projections.

Data availability

The codes supporting the numerical experiments are freely available in the Github repository, https://github.com/maxlemes/SGP.

Bibliography

021inexact [1] A. A. Aguiar, O. P. Ferreira, and L. F. Prudente. Inexact gradient projection method with relative error tolerance. *arXiv preprint arXiv:2101.11146*, 2021.

eandro2020 [2] A. A. Aguiar, O. P. Ferreira, and L. F. Prudente. Subgradient method with feasible inexact projections for constrained convex optimization problems. *Optimization*, 0(0):1–23, 2021, https://doi.org/10.1080/02331934.2021.1902520.

okhosh2012 [3] M. Ahookhosh, K. Amini, and S. Bahrami. A class of nonmonotone Armijo-type line search method for unconstrained optimization. *Optimization*, 61(4):387–404, 2012.

2017linear [4] Z. Allen-Zhu, E. Hazan, W. Hu, and Y. Li. Linear convergence of a Frank-Wolfe type algorithm over trace-norm balls. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, pages 6192–6201, Red Hook, NY, USA, 2017. Curran Associates Inc.

num/drh020 [5] R. Andreani, E. G. Birgin, J. M. Martínez, and J. Yuan. Spectral projected gradient and variable metric methods for optimization with linear inequalities. *IMA J. Numer. Anal.*, 25(2):221–252, 04 2005, https://academic.oup.com/imajna/article-pdf/25/2/221/2090233/drh020.pdf.

rtzner1999 [6] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

Paulo2007 [7] A. Auslender, P. J. S. Silva, and M. Teboulle. Nonmonotone projected gradient methods based on barrier and Euclidean distances. *Comput. Optim. Appl.*, 38(3):305–327, 2007.

BB1988 [8] J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA J. Numer. Anal.*, 8(1):141–148, 1988.

eLivro2014 [9] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer Publishing Company, Incorporated, 1st edition, 2011.

oulle2004 [10] A. Beck and M. Teboulle. A conditional gradient method with linear rate of convergence for solving convex linear systems. *Math. Methods Oper. Res.*, 59(2):235–247, 2004.

roman2010 [11] J. Y. Bello Cruz and L. R. Lucambio Pérez. Convergence of a projected gradient method variant for quasiconvex objectives. *Nonlinear Anal.*, 73(9):2917–2922, 2010.

sekas1976 [12] D. P. Bertsekas. On the Goldstein-Levitin-Polyak gradient projection method. *IEEE Trans. Automat. Control*, 21(2):174–184, 1976.

sekas1999 [13] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific Optimization and Computation Series. Athena Scientific, Belmont, MA, second edition, 1999.

spgsiam [14] E. G. Birgin, J. M. Martínez, and M. Raydan. Nonmonotone spectral projected gradient methods on convex sets. *SIAM J. Optim.*, 10(4):1196–1211, 2000, https://doi.org/10.1137/S1052623497330963.

aydan2003 [15] E. G. Birgin, J. M. Martínez, and M. Raydan. Inexact spectral projected gradient methods on convex sets. *IMA J. Numer. Anal.*, 23(4):539–559, 2003.

SSv060i03 [16] E. G. Birgin, J. M. Martínez, and M. Raydan. Spectral projected gradient methods: Review and perspectives. *J. Stat. Softw.*, 60(3):1–21, 2014.

aydan2005 [17] E. G. Birgin and M. Raydan. Robust stopping criteria for dykstra's algorithm. *SIAM J. Sci. Comput.*, 26(4):1405–1414, 2005.

ttini2016 [18] S. Bonettini, I. Loris, F. Porta, and M. Prato. Variable metric inexact line-search-based methods for nonsmooth optimization. *SIAM J. Optim.*, 26(2):891–921, 2016.

019recent [19] S. Bonettini, F. Porta, M. Prato, S. Rebegoldi, V. Ruggiero, and L. Zanni. Recent advances in variable metric first-order methods. In *Computational Methods for Inverse Problems in Imaging*, pages 1–31. Springer, 2019.

Prato2015 [20] S. Bonettini and M. Prato. New convergence results for the scaled gradient projection method. *Inverse Problems*, 31(9):095008, 20, 2015.

ttini2009 [21] S. Bonettini, R. Zanella, and L. Zanni. A scaled gradient projection method for constrained image deblurring. *Inverse Problems*, 25(1):015002, 23, 2009.

cedal2018 [22] L. Bottou, F. E. Curtis, and J. Nocedal. Optimization methods for large-scale machine learning. *SIAM Rev.*, 60(2):223–311, 2018.

erron1994 [23] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics, 1994, https://epubs.siam.org/doi/pdf/10.1137/1.9781611970777.

kstra1986 [24] J. P. Boyle and R. L. Dykstra. A method for finding projections onto the intersection of convex sets in Hilbert spaces. In *Advances in order restricted statistical inference (Iowa City, Iowa, 1985)*, volume 37 of *Lect. Notes Stat.*, pages 28–47. Springer, Berlin, 1986.

tesVu2013 [25] P. L. Combettes and B. C. Vũ. Variable metric quasi-Fejér monotonicity. *Nonlinear Anal.*, 78:17–31, 2013.

Dai2002 [26] Y. H. Dai. On the nonmonotone line search. *J. Optim. Theory Appl.*, 112(2):315–330, 2002.

tcher2005 [27] Y.-H. Dai and R. Fletcher. Projected Barzilai-Borwein methods for large-scale box-constrained quadratic programming. *Numer. Math.*, 100(1):21–47, 2005.

tcher2006 [28] Y.-H. Dai and R. Fletcher. New algorithms for singly linearly constrained quadratic programs subject to lower and upper bounds. *Math. Program.*, 106(3, Ser. A):403–421, 2006.

iHage2006 [29] Y.-H. Dai, W. W. Hager, K. Schittkowski, and H. Zhang. The cyclic Barzilai-Borwein method for unconstrained optimization. *IMA J. Numer. Anal.*, 26(3):604–627, 2006.

ilson2018 [30] F. R. de Oliveira, O. P. Ferreira, and G. N. Silva. Newton's method with feasible inexact projections for solving constrained generalized equations. *Comput. Optim. Appl.*, 72(1):159–177, 2019.

afino2018 [31] D. di Serafino, V. Ruggiero, G. Toraldo, and L. Zanni. On the steplength selection in gradient methods for unconstrained optimization. *Appl. Math. Comput.*, 318:176–195, 2018.

andro2019 [32] R. Díaz Millán, O. P. Ferreira, and L. F. Prudente. Alternating conditional gradient method for convex feasibility problems. *arXiv e-prints*, page arXiv:1912.04247, Dec 2019, 1912.04247.

chmarking [33] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Math. Program.*, 91(2):201–213, 2002.

kstra1983 [34] R. L. Dykstra. An algorithm for restricted least squares regression. *J. Amer. Statist. Assoc.*, 78(384):837–842, 1983.

g_Yan2019 [35] J. Fan, L. Wang, and A. Yan. An inexact projected gradient method for sparsity-constrained quadratic measurements regression. *Asia-Pac. J. Oper. Res.*, 36(2):1940008, 21, 2019.

verdt2019 [36] N. S. Fazzio and M. L. Schuverdt. Convergence analysis of a nonmonotone projected gradient method for multiobjective optimization problems. *Optim. Lett.*, 13(6):1365–1379, 2019.

iredo2007 [37] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE J. Sel. Topics Signal Process.*, 1(4):586–597, Dec 2007.

ander1999 [38] A. Friedlander, J. M. Martínez, B. Molina, and M. Raydan. Gradient method with retards and generalizations. *SIAM J. Numer. Anal.*, 36(1):275–289, 1999.

llmer2002 [39] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

Hazan2015 [40] D. Garber and E. Hazan. Faster rates for the frank-wolfe method over strongly-convex sets. *Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Volume 37*, pages 541–549, 2015.

avies2018 [41] M. Golbabaee and M. E. Davies. Inexact gradient projection and fast data driven compressed sensing. *IEEE Trans. Inform. Theory*, 64(10):6707–6721, 2018.

stein1964 [42] A. A. Goldstein. Convex programming in Hilbert space. *Bull. Amer. Math. Soc.*, 70:709–710, 1964.

projected [43] D. S. Gonçalves, M. A. Gomes-Ruggiero, and C. Lavor. A projected gradient method for optimization over density matrices. *Optim. Methods Softw.*, 31(2):328–341, 2016, https://doi.org/10.1080/10556788.2015.1082105.

alves2020 [44] D. S. Gonçalves, M. L. N. Gonçalves, and T. C. Menezes. Inexact variable metric method for convex-constrained optimization problems. *Optimization*, 0(0):1–19, 2021, https://doi.org/10.1080/02331934.2021.1887181.

Gong2011 [45] P. Gong, K. Gai, and C. Zhang. Efficient euclidean projections via piecewise root finding and its application in gradient projection. *Neurocomputing*, 74(17):2754 – 2766, 2011.

Sachs2017 [46] G. N. Grapiglia and E. W. Sachs. On the worst-case evaluation complexity of non-monotone line search algorithms. *Comput. Optim. Appl.*, 68(3):555–577, 2017.

Sachs2020 [47] G. N. Grapiglia and E. W. Sachs. A generalized worst-case complexity analysis for non-monotone line searches. *Numer. Algorithms*, 87(2):779–796, Jun 2021.

rippo1986 [48] L. Grippo, F. Lampariello, and S. Lucidi. A nonmonotone line search technique for Newton's method. *SIAM J. Numer. Anal.*, 23(4):707–716, 1986.

igham2002 [49] N. J. Higham. Computing the nearest correlation matrix—a problem from finance. *IMA J. Numer. Anal.*, 22(3):329–343, 2002.

Iusem2003 [50] A. N. Iusem. On the convergence properties of the projected gradient method for convex optimization. *Comput. Appl. Math.*, 22(1):37–52, 2003.

Jaggi2013 [51] M. Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In S. Dasgupta and D. McAllester, editors, *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 427–435, Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR.

vitin1966 [52] E. Levitin and B. Polyak. Constrained minimization methods. *USSR Comput. Math. Math. Phys.*, 6(5):1 – 50, 1966.

u_Gao2015 [53] G. Ma, Y. Hu, and H. Gao. An accelerated momentum based gradient projection method for image deblurring. In *2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, pages 1–4, 2015.

owitz1952 [54] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

iuYan2007 [55] J. Mo, C. Liu, and S. Yan. A nonmonotone trust region method based on nonincreasing technique of weighted average of the successive function values. *J. Comput. Appl. Math.*, 209(1):97–108, 2007.

More1990 [56] J. J. Moré. On the performance of algorithms for large-scale bound constrained problems. In *Large-scale numerical optimization (Ithaca, NY, 1989)*, pages 32–45. SIAM, Philadelphia, PA, 1990.

ovski2013 [57] Y. Nesterov and A. Nemirovski. On first-order algorithms for ℓ_1/nuclear norm minimization. *Acta Numer.*, 22:509–575, 2013.

numerical [58] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

anier1991 [59] E. R. Panier and A. L. Tits. Avoiding the Maratos effect by means of a nonmonotone line search. I. General constrained problems. *SIAM J. Numer. Anal.*, 28(4):1183–1195, 1991.

coara2018 [60] A. Patrascu and I. Necoara. On the convergence of inexact projection primal first-order methods for convex minimization. *IEEE Trans. Automat. Control*, 63(10):3317–3329, 2018.

Livro1987 [61] B. T. Polyak. *Introduction to optimization*. Translations Series in Mathematics and Engineering. Optimization Software, Inc., Publications Division, New York, 1987. Translated from the Russian, With a foreword by Dimitri P. Bertsekas.

Rasch2020 [62] J. Rasch and A. Chambolle. Inexact first-order primal-dual algorithms. *Comput. Optim. Appl.*, 76(2):381–430, 2020.

ykstraSDD [63] M. Raydan and P. Tarazaga. Primal and polar approach for computing the symmetric diagonally dominant projection. *Numer. Linear Algebra Appl.*, 9(5):333–345, 2002, https://onlinelibrary.wiley.com/doi/pdf/10.1002/nla.277.

yasev2001 [64] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, (2):21–41, 2001.

Sachs2011 [65] E. W. Sachs and S. M. Sachs. Nonmonotone line searches for optimization algorithms. *Control Cybernet.*, 40(4):1059–1075, 2011.

Villa2012 [66] S. Salzo and S. Villa. Inexact and accelerated proximal point algorithms. *J. Convex Anal.*, 19(4):1167–1192, 2012.

Spinu2013 [67] F. Spinu. An algorithm for computing risk parity weights. *SSRN Electronic Journal*, 2013.

right2012 [68] S. Sra, S. Nowozin, and S. Wright. *Optimization for Machine Learning*. Neural information processing series. MIT Press, 2012.

avies2017 [69] J. Tang, M. Golbabaee, and M. E. Davies. Gradient projection iterative sketch for large-scale constrained least-squares. In D. Precup and Y. W. Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 3377–3386, International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR.

Toint1996 [70] P. L. Toint. An assessment of nonmonotone linesearch techniques for unconstrained optimization. *SIAM J. Sci. Comput.*, 17(3):725–739, 1996.

Salzo2013 [71] S. Villa, S. Salzo, L. Baldassarre, and A. Verri. Accelerated and inexact forward-backward algorithms. *SIAM J. Optim.*, 23(3):1607–1633, 2013.

ngLiu2005 [72] C. Wang, Q. Liu, and X. Yang. Convergence properties of nonmonotone spectral projected gradient methods. *J. Comput. Appl. Math.*, 182(1):51–66, 2005.

ihong2018 [73] X. Yan, K. Wang, and H. He. On the convergence rate of scaled gradient projection method. *Optimization*, 67(9):1365–1376, 2018.

_Yang2019 [74] F. Zhang, H. Wang, J. Wang, and K. Yang. Inexact primal–dual gradient projection methods for nonlinear optimization on convex set. *Optimization*, 69(10):2339–2365, 2020, https://doi.org/10.1080/02331934.2019.1696338.

Hager2004 [75] H. Zhang and W. W. Hager. A nonmonotone line search technique and its application to unconstrained optimization. *SIAM J. Optim.*, 14(4):1043–1056, 2004.

Dai2006 [76] B. Zhou, L. Gao, and Y.-H. Dai. Gradient methods with adaptive step-sizes. *Comput. Optim. Appl.*, 35(1):69–86, 2006.